

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

57414

From: Nickol, Gary
Sent: Wednesday, January 02, 2002 2:03 PM
To: STIC-Biotech/ChemLib
Subject: 09/733956

1) Please search the amino acids of SEQ ID NO:2.

2) Please search DNA encoding SEQ ID NO:2.

Thanks,

Gary Nickol Ph.D.
AU:1642, Room 8D09, Mailbox 8E12
703-305-7143

CRFE

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 1/9/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:16 ; Search time 12.66 Seconds
(without alignments)
719.892 Million cell updates/sec

Title: US-09-733-956-2

Perfect score: 2091
Sequence: 1 MNSTESNSASVREKVLSTI.....SQEASQTLDSVSHLPDL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sequences: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2091	100.0	655	3	US-08-959-382-2
2	1914	91.5	573	4	US-09-042-785A-2
3	1527	73.0	605	4	US-09-042-785A-23
4	431	20.6	84	4	US-09-042-785A-30
5	431	20.6	84	4	US-09-042-785A-31
6	115.5	5.5	2183	3	US-08-746-111-5
7	99	4.7	2343	4	US-09-324-867-2
8	95.5	4.6	287	4	US-08-985-950-6
9	95.5	4.6	287	4	US-08-985-950-8
10	93.5	4.4	465	3	US-08-788-231A-2
11	91.5	4.4	1093	4	US-09-315-793-52
12	90	4.3	1719	2	US-08-459-568-4
13	90	4.3	1719	2	US-08-399-411-4
14	90	4.3	1719	3	US-08-516-859A-4
15	89	4.3	761	4	US-08-928-941D-1
16	89	4.3	761	4	US-09-280-590A-1
17	87.5	4.2	760	4	US-08-928-941D-29
18	87.5	4.2	760	4	US-09-280-590A-29
19	87	4.2	651	4	US-08-985-950-22
20	86.5	4.1	895	4	US-08-827-962-19
21	86.5	4.1	895	4	US-08-827-962-21
22	86.5	4.1	1072	4	US-09-357-251-31
23	86.5	4.1	1162	4	US-08-827-962-15
24	86.5	4.1	1162	4	US-08-827-962-20
25	86.5	4.1	1162	4	US-08-803-346-1
26	86	4.1	1248	2	US-09-080-897-2
27	86	4.1	1248	4	US-09-323-735-2

28	86	4.1	1315	3	US-08-899-595-3	Sequence 3, Appl
29	85.5	4.1	339	1	US-08-431-080-18	Sequence 18, Appl
30	85.5	4.1	339	2	US-08-938-53A-18	Sequence 18, Appl
31	85.5	4.1	1447	4	US-09-041-886-25	Sequence 25, Appl
32	85.5	4.1	1447	5	PCR-US94-05277-2	Sequence 2, Appl
33	85	4.1	405	4	US-09-413-574-2	Sequence 2, Appl
34	84.5	4.0	599	4	US-09-000-145-2	Sequence 2, Appl
35	84.5	4.0	646	4	US-09-625-188-10	Sequence 10, Appl
36	84	4.0	772	1	US-08-524-757-12	Sequence 12, Appl
37	84	4.0	2548	4	US-09-172-422-1	Sequence 1, Appl
38	83.5	4.0	1070	4	US-08-697-954-2	Sequence 1, Appl
39	83	4.0	1481	2	US-08-616-844-40	Sequence 40, Appl
40	83	4.0	1481	2	US-08-599-654-40	Sequence 40, Appl
41	83	4.0	1481	3	US-08-944-868A-40	Sequence 40, Appl
42	83	4.0	1481	3	US-08-944-422A-40	Sequence 40, Appl
43	83	4.0	1481	3	US-08-944-496-40	Sequence 40, Appl
44	82.5	3.9	771	3	US-09-121-964-9	Sequence 9, Appl
45	82.5	3.9	1230	2	US-08-968-542C-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-08-959-382-2
Sequence 2, Application US/08959382
Patent No. 6013476
GENERAL INFORMATION:
APPLICANT: DEEN, KEITH
APPLICANT: HURLE, MARK
APPLICANT: YOUNG, PETER
APPLICANT: TAN, K. B.
TITLE OF INVENTION: Tumor Necrosis Related Receptor,
TITLE OF INVENTION: TR7
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P. O. BOX 980
CITY: VALLEY Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,382
FILING DATE: 28-OCT-1997
CLASSIFICATION: 544
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/041,796
FILING DATE: 02-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-50017
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 655 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-959-382-2
Query Match 100.0%; Score 2091; DB 3; Length 655;

THIS PAGE BLANK (USPTO)

Best Local Similarity 100.0%; Pred. No. 4.6e-204.
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTLPNLQVNNHQQGPHRHIL 60
DB 251 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTLPNLQVNNHQQGPHRHIL 310
QY 61 KLLPSMATGGEKSTPIKPKRGPHRONLHKHFDINEHLPMMIVLFLVLLVIVVCS 120
DB 311 KLLPSMATGGEKSTPIKPKRGPHRONLHKHFDINEHLPMMIVLFLVLLVIVVCS 370
QY 121 RKSSRTLKKGPDPDPAIVKAGLKKSMPTONREKWIYCCNGHIDILKLVAAQVGSQ 180
DB 371 RKSSRTLKKGPDPDPAIVKAGLKKSMPTONREKWIYCCNGHIDILKLVAAQVGSQ 430
QY 181 KDIOFLCNASEREVAFAFNGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 240
DB 431 KDIOFLCNASEREVAFAFNGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 490
QY 241 VERIRGLMEDTQLETDLKALPMSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 300
DB 491 VERIRGLMEDTQLETDLKALPMSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 550
QY 301 VDSEPLRLCDSTSSGSSALSRNGSFITREKKDVTLRQVRLDPCDLPIDDMHLFNP 360
DB 551 VDSEPLRLCDSTSSGSSALSRNGSFITREKKDVTLRQVRLDPCDLPIDDMHLFNP 610
QY 361 ELRVIEEIPQAEKDLRLFEIIGVKSQASQTLDSVYSHLPDLL 405
DB 611 ELRVIEEIPQAEKDLRLFEIIGVKSQASQTLDSVYSHLPDLL 655

RESULT 2
US-09-042-785A-2
Sequence 2, Application US/09042785A

Patent No. 6194151
GENERAL INFORMATION:
APPLICANT: Busfield, Samantha J
TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
NUMBER OF INVENTIONS: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
MEDICAL READABLE FORM:
COMPUTER TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,785A
FILING DATE: 17-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/938,896
FILING DATE: 26-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MEI-001CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-4214
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 573 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-785A-2

Query Match 91.5%; Score 1914; DB 4; Length 573;
Best Local Similarity 92.9%; Pred. No. 3.8e-186.
Matches 377; Conservative 6; Mismatches 21; Indels 2; Gaps 2;

QY 1 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTLPNLQVNNHQQGPHRHIL 60
DB 169 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTLPNLQVNNHQQGPHRHIL 228
QY 61 KLLP-SMATGGEKSTPIKPKRGPHRONLHKHFDINEHLPMMIVLFLVLLVIVVCS 119
DB 229 KLLP-SMATGGEKSTPIKPKRGPHRONLHKHFDINEHLPMMIVLFLVLLVIVVCS 287
QY 120 IRKSSRTLKKGPDPDPAIVKAGLKKSMPTONREKWIYCCNGHIDILKLVAAQVGSQ 179
DB 288 IRKSSRTLKKGPDPDPAIVKAGLKKSMPTONREKWIYCCNGHIDILKLVAAQVGSQ 347
QY 180 WKDIYFLCNASEREVAFAFNGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 239
DB 348 WKDIYFLCNASEREVAFAFNGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 407
QY 240 VVERIRGLMEDTQLETDLKALPMSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 299
DB 408 VVERIRGLMEDTQLETDLKALPMSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 467
QY 300 VDSEPLRLCDSTSSGSSALSRNGSFITREKKDVTLRQVRLDPCDLPIDDMHLFNP 359
DB 468 VDSEPLRLCDSTSSGSSALSRNGSFITREKKDVTLRQVRLDPCDLPIDDMHLFNP 527
QY 360 ELRVIEEIPQAEKDLRLFEIIGVKSQASQTLDSVYSHLPDLL 405
DB 528 ELRVIEEIPQAEKDLRLFEIIGVKSQASQTLDSVYSHLPDLL 573

RESULT 3
US-09-042-785A-23
Sequence 23, Application US/09042785A

Patent No. 6194151
GENERAL INFORMATION:
APPLICANT: Busfield, Samantha J
TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
NUMBER OF INVENTIONS: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
MEDICAL READABLE FORM:
COMPUTER TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,785A
FILING DATE: 17-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/938,896
FILING DATE: 26-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MEI-001CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-4214
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:16 ; Search time 12.66 seconds

(without alignments)
719.892 Million cell updates/sec

Title: US-09-733-956-2

Perfect score: 2091

Sequence: 1 MNSTESNSASVRRPKVLSSI.....SQEASQRLDSVSHLPDLL 405

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6CTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2091	100.0	655	3	US-08-959-382-2
2	1914	91.5	573	4	US-09-042-785A-2
3	1527	73.0	605	4	US-09-042-785A-23
4	431	20.6	84	4	US-09-042-785A-30
5	431	20.6	84	4	US-09-042-785A-31
6	115.5	5.5	2183	3	US-08-746-111-5
7	99	4.7	2343	4	US-09-324-867-2
8	95.5	4.6	287	4	US-08-985-950-6
9	95.5	4.6	287	4	US-08-985-950-8
10	93	4.4	465	3	US-08-788-231A-2
11	91.5	4.4	1093	4	US-09-315-793-52
12	90	4.3	1719	2	US-08-459-568-4
13	90	4.3	1719	2	US-08-399-411-4
14	90	4.3	1719	3	US-08-516-859A-4
15	89	4.3	761	4	US-08-928-941D-1
16	89	4.3	761	4	US-09-280-590A-1
17	87.5	4.2	760	4	US-08-928-941D-29
18	87.5	4.2	760	4	US-09-280-590A-29
19	87	4.2	651	4	US-08-985-950-22
20	86.5	4.1	895	4	US-08-827-962-19
21	86.5	4.1	895	4	US-08-827-962-21
22	86.5	4.1	1072	4	US-09-357-251-31
23	86.5	4.1	1162	4	US-08-827-962-15
24	86.5	4.1	1162	4	US-08-827-962-20
25	86.5	4.1	1162	4	US-08-803-346-1
26	86	4.1	1248	2	US-09-080-897-2
27	86	4.1	1248	4	US-09-323-735-2

28	86	4.1	1315	3	US-08-899-595-3	Sequence 3, App1
29	85.5	4.1	339	1	US-08-431-080-18	Sequence 18, App1
30	85.5	4.1	339	2	US-08-938-534-18	Sequence 18, App1
31	85.5	4.1	1447	5	US-09-041-886-25	Sequence 25, App1
32	85.5	4.1	1447	5	PCR-US94-05277-2	Sequence 2, App1
33	85	4.1	405	4	US-09-413-574-2	Sequence 2, App1
34	84.5	4.0	599	4	US-09-000-145-2	Sequence 2, App1
35	84.5	4.0	646	4	US-09-625-188-10	Sequence 10, App1
36	84	4.0	772	1	US-08-524-757-12	Sequence 12, App1
37	84	4.0	2548	4	US-09-172-422-1	Sequence 1, App1
38	83.5	4.0	1070	4	US-08-697-954-2	Sequence 2, App1
39	83	4.0	1481	2	US-08-616-844-40	Sequence 40, App1
40	83	4.0	1481	2	US-08-599-654-40	Sequence 40, App1
41	83	4.0	1481	3	US-08-944-868A-40	Sequence 40, App1
42	83	4.0	1481	3	US-08-944-423A-40	Sequence 40, App1
43	83	4.0	1481	3	US-08-944-496-40	Sequence 40, App1
44	82.5	3.9	771	3	US-09-121-964-9	Sequence 9, App1
45	82.5	3.9	1230	2	US-08-968-542C-35	Sequence 35, App1

ALIGNMENTS

RESULT 1
US-08-959-382-2
; Sequence 2, Application US/08959382
; Patent No. 6013476
; GENERAL INFORMATION:
; APPLICANT: DEEN, KEITH
; APPLICANT: HURLE, MARK
; APPLICANT: YOUNG, PETER
; APPLICANT: TAN, K. B.
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P. O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,382
; FILING DATE: 28-OCT-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,796
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F.
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 655 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-959-382-2
Query Match 100.0%; Score 2091; DB 3; length 655;

Best Local Similarity 100.0%; Pred. No. 4.6e-204;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MNSTSNSSASVRKRVLSSTIOEGTVPDNTSSARGKEDVNTKLPNLQVYVNHQOGPHHRIIL 60
Db 251 MNSTSNSSASVRKRVLSSTIOEGTVPDNTSSARGKEDVNTKLPNLQVYVNHQOGPHHRIIL 310
QY 61 KLPSMEATGEGKSTPIKGRGHRPNLHKHFDINELPMIVLFLLVLVVYVCS 120
Db 311 KLPSMEATGEGKSTPIKGRGHRPNLHKHFDINELPMIVLFLLVLVVYVCS 370
QY 121 RKSRTLKGGKGRPPSAIVEKAGLKSMTPTQNRKRWIYNGHGIDILKLVAAQVGSQ 180
Db 371 RKSRTLKGGKGRPPSAIVEKAGLKSMTPTQNRKRWIYNGHGIDILKLVAAQVGSQ 430
QY 181 KDIYQFLCNASEREVAAFNSGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 240
Db 431 KDIYQFLCNASEREVAAFNSGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 490
QY 241 VEKIRGLMEDTTOLETDKLALPMSPLSPSPISPNAKLENSALLTVESPQDNKGF 300
Db 491 VEKIRGLMEDTTOLETDKLALPMSPLSPSPISPNAKLENSALLTVESPQDNKGF 550
QY 301 VDESEPLRCSTSSGSSALSRNGSFITKEKKDVTLRQVRLDPCDLPIDFDMHLFNP 360
Db 551 VDESEPLRCSTSSGSSALSRNGSFITKEKKDVTLRQVRLDPCDLPIDFDMHLFNP 610
QY 361 ELRYIEEIPQAEKLDRLFEIIGVKSQESQTLDSVYSHLPDLL 405
Db 611 ELRYIEEIPQAEKLDRLFEIIGVKSQESQTLDSVYSHLPDLL 655
```

RESULT 2
US-09-042-785A-2

; Sequence 2, Application US/09042785A
; Patent No. 6194151
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J
; TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042.785A
; FILING DATE: 17-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/938,896
; FILING DATE: 26-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MEI-001CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 573 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-042-785A-2

Query Match 91.5%; Score 1914; DB 4; Length 573;

Best Local Similarity 92.9%; Pred. No. 3.8e-186;
Matches 377; Conservative 6; Mismatches 21; Indels 2; Gaps 2;

```
QY 1 MNSTSNSSASVRKRVLSSTIOEGTVPDNTSSARGKEDVNTKLPNLQVYVNHQOGPHHRIIL 60
Db 169 MNSTSNSSASVRKRVLSSTIOEGTVPDNTSSARGKEDVNTKLPNLQVYVNHQOGPHHRIIL 228
QY 61 KLPSMEATGEGKSTPIKGRGHRPNLHKHFDINELPMIVLFLLVLVVYVCS 119
Db 229 KLPSMEATGEGKSTPIKGRGHRPNLHKHFDINELPMIVLFLLVLVVYVCS 287
QY 120 IRKSRTLKGGKGRPPSAIVEKAGLKSMTPTQNRKRWIYNGHGIDILKLVAAQVGSQ 179
Db 288 IRKSRTLKGGKGRPPSAIVEKAGLKSMTPTQNRKRWIYNGHGIDILKLVAAQVGSQ 347
QY 180 KDIYQFLCNASEREVAAFNSGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 239
Db 348 KDIYQFLCNASEREVAAFNSGYTADHERAYALQHWITRGPEASLAQLISALRQHRND 407
QY 240 VEKIRGLMEDTTOLETDKLALPMSPLSPSPISPNAKLENSALLTVESPQDNKGF 299
Db 408 VEKIRGLMEDTTOLETDKLALPMSPLSPSPISPNAKLENSALLTVESPQDNKGF 467
QY 300 VDESEPLRCSTSSGSSALSRNGSFITKEKKDVTLRQVRLDPCDLPIDFDMHLFNP 359
Db 468 VDESEPLRCSTSSGSSALSRNGSFITKEKKDVTLRQVRLDPCDLPIDFDMHLFNP 527
QY 360 ELRYIEEIPQAEKLDRLFEIIGVKSQESQTLDSVYSHLPDLL 405
Db 528 ELRYIEEIPQAEKLDRLFEIIGVKSQESQTLDSVYSHLPDLL 573
```

RESULT 3
US-09-042-785A-23

; Sequence 23, Application US/09042785A
; Patent No. 6194151
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J
; TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042.785A
; FILING DATE: 17-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/938,896
; FILING DATE: 26-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MEI-001CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-042-785A-23

Query Match 73.0%; Score 1527; DB 4; Length 605;
Best Local Similarity 100.0%; Pred. No. 1e-146;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNSTESNSASVPRKYLSTIOECTVPDNMTSSAGKEDVNTLNLQVNNHQQGPHRHIL 60
DB 251 MNSTESNSASVPRKYLSTIOECTVPDNMTSSAGKEDVNTLNLQVNNHQQGPHRHIL 310
OY 61 KLPSMEATGGEKSTPIKGRGHRONLHKHFDINELHPMIVFLVLVYVVCST 120
DB 311 KLPSMEATGGEKSTPIKGRGHRONLHKHFDINELHPMIVFLVLVYVVCST 370
OY 121 RKSSRTLKGRPDPSAIVERKAGLKKSMPTONREKWIYVNCNGHIDILKLVAAVGSQW 180
DB 371 RKSSRTLKGRPDPSAIVERKAGLKKSMPTONREKWIYVNCNGHIDILKLVAAVGSQW 430
OY 181 KDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPESLAQLISALRQHRNDV 240
DB 431 KDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPESLAQLISALRQHRNDV 490
OY 241 VERIRGLMEDTQLETDKLALPMSPLSPSPISPAPAKLNSALLTVESPDP 294
DB 491 VERIRGLMEDTQLETDKLALPMSPLSPSPISPAPAKLNSALLTVESPDP 544

RESULT 4

US-09-042-785A-30
Sequence 30, Application US/09042785A

Patent No. 6194151

GENERAL INFORMATION:

APPLICANT: Busfield, Samantha J
TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY

TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,785A

FILING DATE: 17-MAR-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/938,896

FILING DATE: 26-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MEI-001CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: Internal

US-09-042-785A-30

Query Match 20.6%; Score 431; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 165 GIDILKLVAAVGSQWMDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPES 224
DB 1 GIDILKLVAAVGSQWMDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPES 60
OY 225 LAQLISALRQHRNDVYEKIRGLM 248
DB 61 LAQLISALRQHRNDVYEKIRGLM 84

RESULT 5

US-09-042-785A-31
Sequence 31, Application US/09042785A

Patent No. 6194151

GENERAL INFORMATION:

APPLICANT: Busfield, Samantha J
TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY

TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,785A

FILING DATE: 17-MAR-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/938,896

FILING DATE: 26-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MEI-001CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: Internal

US-09-042-785A-31

Query Match 20.6%; Score 431; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 165 GIDILKLVAAVGSQWMDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPES 224
DB 1 GIDILKLVAAVGSQWMDIYQFLCNASEREVAAFSGYTADEHRAVAALQHWITIRGPES 60
OY 225 LAQLISALRQHRNDVYEKIRGLM 248
DB 61 LAQLISALRQHRNDVYEKIRGLM 84

RESULT 6

US-08-746-111-5


```

US-08-985-950-6
; Sequence 6, Application US/08985950-
; Patent No. 6140076
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,950
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,279
; FILING DATE: 21-MARCH-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,181
; FILING DATE: 16-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,252
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0670K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1204
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-985-950-6

```

```

Query Match 4.6%; Score 95.5; DB 4; Length 287;
Best Local Similarity 21.6%; Pred. No. 0.16;
Matches 50; Conservative 32; Mismatches 75; Indels 75; Gaps 12;

QY 27 DNTSSARKEDVNTKLPNL-----QVYVNHQGFH-----HRIHLKLLPSM 66
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 ESRSTYNDTEDEVSQSPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSQSDYLELVK- 121
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 EATGGEKSTPIKPKRG---HPRQNLHKHFDINEHLP-----WMIVLFL 109
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 122 ETSGGPDSPTDEGSSAGTQRPDSNSH-----NEHAPASQGLKAHLYLLGVSVPFL 176
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 110 LVLVVIVCSIRKSSRTLKGGPRDPSAIVERAGLKKSMPTPNRREKWTYYCNGHSIDL 169
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 CLLLVFLCLHRQNO--IKQGPFRSKD-----EQQKPOORPDL-----AVDYL 217
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 KLVAAOV---GSQMKDIYQFLCNASEREVAAPFNGYTADHERAYALQHWHTI 218
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 ERTADKATVNGLPEDK-----RETDTSALAG--SSQEVYTYAOLDHWAL 259
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
US-08-985-950-8
; Sequence 8, Application US/08985950
; Patent No. 6140076

```

```

; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,950
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,279
; FILING DATE: 21-MARCH-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,181
; FILING DATE: 16-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,252
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0670K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1204
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-985-950-8

```

```

Query Match 4.6%; Score 95.5; DB 4; Length 287;
Best Local Similarity 21.6%; Pred. No. 0.16;
Matches 50; Conservative 32; Mismatches 75; Indels 75; Gaps 12;

QY 27 DNTSSARKEDVNTKLPNL-----QVYVNHQGFH-----HRIHLKLLPSM 66
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 ESRSTYNDTEDEVSQSPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSQSDYLELVK- 121
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 67 EATGGEKSTPIKPKRG---HPRQNLHKHFDINEHLP-----WMIVLFL 109
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 122 ETSGGPDSPTDEGSSAGTQRPDSNSH-----NEHAPASQGLKAHLYLLGVSVPFL 176
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 110 LVLVVIVCSIRKSSRTLKGGPRDPSAIVERAGLKKSMPTPNRREKWTYYCNGHSIDL 169
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 CLLLVFLCLHRQNO--IKQGPFRSKD-----EQQKPOORPDL-----AVDYL 217
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 KLVAAOV---GSQMKDIYQFLCNASEREVAAPFNGYTADHERAYALQHWHTI 218
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 ERTADKATVNGLPEDK-----RETDTSALAG--SSQEVYTYAOLDHWAL 259
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 10
US-08-788-231A-2
; Sequence 2, Application US/08788231A
; Patent No. 6019974
; GENERAL INFORMATION:
; APPLICANT: L'Hernault, Steven W.
; TITLE OF INVENTION: SPE-4 RELATED PEPTIDES, ANTIBODIES AND

```

TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,231A
FILING DATE: 24-JAN-1997
CLASSIFICATION: 42A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010,672
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Feiber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 60-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8080
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
US-08-788-231A-2

Query Match 4.4%; Score 93; DB 3; Length 465;
Best Local Similarity 23.7%; Pred. No. 0.65;
Matches 47; Conservative 33; Mismatches 62; Indels 56; Gaps 10;

QY 169 LKVAAGQSMKDYQLC-NASEREVAAPNGYTDHERAYALQHTTIRGPASLAQ 227
||| : : : : :
DB 211 LKKVQEKASDYKCVLNLIMESANEKRLTAGSN-----QETWEGESTIR 257
||| : : : : :
QY 228 LIS-----AIRQHRNRYVEKIRGIMEDTQLETDKLALPMSPLSPSPSPNAKLE 281
||| : : : : :
DB 258 TVKQTIERYTKREADDEFYQIR-----ORRAINPDSVPT-----E 295
||| : : : : :
QY 282 NSALITVEPSP---QDKN-KGFVDESEPLRLCDSTSSGSSALSRSNGSFITKEKKDTVL 336
||| : : : : :
DB 296 HSPIVEAPSPSELEKKEKSTELSDSD---TSETSSGSSNLSSSSDSTTVTSRSDI-- 349
||| : : : : :
QY 337 ROYRLDPCDLPDPDML 354
||| : : : : :
DB 350 --STAECD-OKEMDLY 364
||| : : : : :

RESULT 11
US-09-315-793-52
Sequence 52, Application US/09345793
Patent No. 6221597
GENERAL INFORMATION:
APPLICANT: Roberts, Christopher J.
TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
FILE REFERENCE: 9301-048
CURRENT APPLICATION NUMBER: US/09/315,793
CURRENT FILING DATE: 1999-05-21;
NUMBER OF SEQ ID NOS: 62

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1093
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-315-793-52

Query Match 4.4%; Score 91.5; DB 4; Length 1093;
Best Local Similarity 22.4%; Pred. No. 3.7;
Matches 73; Conservative 51; Mismatches 119; Indels 83; Gaps 17;

QY 142 AGLKSMPTPTONREKMYICGCHGIDILKLAAYGVS-----QMKITYQ--FLCN 189
||| : : : : :
DB 595 AQIKRLIPRPN-GKLEKRIIHGRVLDIKQSAVGSQVPTDVSIKQTFYQSGISMN 653
||| : : : : :
QY 190 AS-----EREVAAFNSGY-----TAD-----HERAYALQH----- 215
||| : : : : :
DB 654 EOKIRIENEITLNKNEYNDKRSITDALSNOKSGYRHELSLASKNNDINREAHQINERK 713
||| : : : : :
QY 216 -WTR-GPEASIALQILSALRRNDVYEKIRGIMEDTQLETDKLALPMSPLSPSP 273
||| : : : : :
DB 714 KYTRKRSITETLRKLDQKREARKDVQKIKDIDQIQQLLKQRHL-LSKMASSMKSL 772
||| : : : : :
QY 274 PSPNAKLENSALLVEPSPQDKN-----GFVDESEPLRLCDSTSSGSSALSRSNGSFITK 329
||| : : : : :
DB 773 KNCQELISTQILOFEAGNMDSVNDYIGFNERADL-----KSQYEDKKKEV-K 822
||| : : : : :
QY 330 EKKDTV-----LROYRLDPCDLPDIFD-DMLHPLNPELRYIEE---IPQAEKDLRL- 378
||| : : : : :
DB 823 EMRDTPEFQSWRREIR-----SYDDTKKLLKNAVEKIEEGNLFSLVQVDYDLKLE 874
||| : : : : :
QY 379 FEIGVKSQEAQOTLLDSVYSHLPDL 404
||| : : : : :
DB 875 SEIAVNHDESAVITLDQVTAELREL 900
||| : : : : :

RESULT 12
US-08-459-568-4
Sequence 4, Application US/08459568
Patent No. 5811304
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,568
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-568-4

```

```

Query Match          4.3%; Score 90; DB 2; Length 1719;
Best Local Similarity 25.4%; Pred. No. 11;
Matches 35; Conservative 15; Mismatches 50; Indels 38; Gaps 5;

```

```

QY 261 LPMSPSPSPSPSPSP--NAKLNSALL--TVEPSPODKNKGFEVDESEPLLRCS--- 312
   || || || || || || || || || || || || || || || || || || || ||
DB 996 LPAPSSASPHPCPSPLSNATQSPILSPYSPSPSP-----IPVEPLMSASAPGP 1049
QY 313 --TSSGSSALSRNGSFITKEKDTVLQVRLDPCDLPIDFDMHLNPELRLVIEIRQ 370
   || || || || || || || || || || || || || || || || || || || ||
DB 1050 PTLSSSSSSSSSSSSSSSSSSSSSSSP-----SPPPLSAISSVVS 1087
QY 371 AEDKLDRLFEIIGVKSQE 388
DB 1088 SGNLEASLPMISFKOE 1105

```

RESULT 13

```

US-08-399-411-4
; Sequence 4, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-399-411-4

```

```

Query Match          4.3%; Score 90; DB 2; Length 1719;
Best Local Similarity 25.4%; Pred. No. 11;
Matches 35; Conservative 15; Mismatches 50; Indels 38; Gaps 5;

```

```

QY 261 LPMSPSPSPSPSP--NAKLNSALL--TVEPSPODKNKGFEVDESEPLLRCS--- 312
   || || || || || || || || || || || || || || || || || || || ||
DB 996 LPAPSSASPHPCPSPLSNATQSPILSPYSPSPSP-----IPVEPLMSASAPGP 1049

```

```

QY 313 --TSSGSSALSRNGSFITKEKDTVLQVRLDPCDLPIDFDMHLNPELRLVIEIRQ 370
   || || || || || || || || || || || || || || || || || || || ||
DB 1050 PTLSSSSSSSSSSSSSSSSSSSSSSSP-----SPPPLSAISSVVS 1087
QY 371 AEDKLDRLFEIIGVKSQE 388
DB 1088 SGNLEASLPMISFKOE 1105

```

RESULT 14

```

US-08-516-859A-4
; Sequence 4, Application US/08516859A
; Patent No. 6069231
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516,859A
; FILING DATE: 18-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-516-859A-4

```

```

Query Match          4.3%; Score 90; DB 3; Length 1719;
Best Local Similarity 25.4%; Pred. No. 11;
Matches 35; Conservative 15; Mismatches 50; Indels 38; Gaps 5;

```

```

QY 261 LPMSPSPSPSPSP--NAKLNSALL--TVEPSPODKNKGFEVDESEPLLRCS--- 312
   || || || || || || || || || || || || || || || || || || || ||
DB 996 LPAPSSASPHPCPSPLSNATQSPILSPYSPSPSP-----IPVEPLMSASAPGP 1049
QY 313 --TSSGSSALSRNGSFITKEKDTVLQVRLDPCDLPIDFDMHLNPELRLVIEIRQ 370
   || || || || || || || || || || || || || || || || || || || ||
DB 1050 PTLSSSSSSSSSSSSSSSSSSSSSSSP-----SPPPLSAISSVVS 1087
QY 371 AEDKLDRLFEIIGVKSQE 388
DB 1088 SGNLEASLPMISFKOE 1105

```

RESULT 15
US-08-928-941D-1
; Sequence 1, Application US/08928941D
; Patent No. 6180763
; GENERAL INFORMATION:
; APPLICANT: Hitai, Hiroshi
; APPLICANT: Sheri, Charles
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,941D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 761 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; US-08-928-941D-1

Query Match 4.3%; Score 89; DB 4; Length 761;

Best Local Similarity 20.6%; Pred. No. 3.7;
Matches 72; Conservative 38; Mismatches 113; Indels 126; Gaps 17;

QY 136 SAIVKAGLKKSMPTQWREKRIYVCGHGDILKLVAAVGSOW--KDIYQFLCNASER 193
DB 307 AAVAEVGVTR--SEKQCRSKWLNLYNM-----KSGGTWTKEDENLILRIAE 354
QY 194 EVA-----AFSNGYADHERAYALQHWYIR-----GPEASLAQLISALRQHRND- 239
DB 355 DVADENDINWDLAEGWSSVRSPOGLRSKMWYIKQIANHKKDVSFVLIKGLKQLENOK 414
QY 240 ---VKEIRG-----LMEDTQLETDKLALPMSPPSPSPPIP- 274
DB 415 NNPLYLENKSGSGVPSNCSNVQHVQIRVARLEDNTAI-----SPSPMALQIPV 465
QY 275 -----SPNAK-----LENSALLTVE-----PSPQDNKGFFVDESEPLLRCDST 313
DB 466 QITHVSSDTSPPAASADSETITLNSGTLOFFELPSFPIQPTGTPETV-----LIQTSSS 519
QY 314 SSGSSALSRLNGSFYTKERKKDYLRQVRRLDPCDLQPIFDDMLHFLNPEEL----- 362
DB 520 QGLPLTLITN-----PTLTIAAAPASPEQI---IYHALSPHLLNTSDNVTVQC 566

QY 363 ---RVIEEIPQAEKLDRLFEIIGVKSOE---ASQTLDSVYSHLPULL 405
DB 567 HTPRVIIQTVATED-----ITSSLQSEELTVDSLDLHSDPFPEPPDAL 608

Search completed: January 4, 2002, 08:38:13
Job time: 117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:11 : Search time 24.92 Seconds
(without alignments)
1203.840 Million cell updates/sec

Title: US-09-733-956-2

Perfect score: 2091
Sequence: 1 MNSTESNSASVAPKVLSSI.....SQEASQTLDSVYSHLPDLL 405

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2091	100.0	405	21	AAV67947
2	2091	100.0	631	20	AAV05678
3	2091	100.0	635	19	AAW5792
4	2091	100.0	655	20	AAV41693
5	2091	100.0	655	20	AAW81059
6	2091	100.0	655	21	AAW26981
7	2091	100.0	655	21	AAW01348
8	2091	100.0	655	21	AAW01338
9	2091	100.0	655	21	AAV77460
10	2091	100.0	691	21	AAV92846
11	2089	99.9	655	21	AAW44249

12	2087	99.8	655	22	AAW93023
13	1914	91.5	573	20	AAV05695
14	1527	73.0	605	20	AAV05697
15	1396	66.8	600	21	AAV92845
16	1345	64.3	393	20	AAV28450
17	1322	63.2	508	20	AAV05680
18	868	41.5	444	21	AAV77461
19	439	21.0	85	22	AAW15484
20	439	21.0	85	22	AAW15484
21	439	21.0	85	22	AAW27972
22	439	21.0	85	22	AAW29434
23	439	21.0	85	22	AAW03241
24	439	21.0	85	22	AAW04642
25	431	20.6	84	20	AAV05698
26	431	20.6	84	20	AAV05699
27	397	19.0	76	22	AAW17149
28	397	19.0	76	22	AAW29641
29	397	19.0	76	22	AAW04844
30	352	16.8	67	21	AAW26987
31	352	16.8	67	21	AAW57452
32	119	5.7	2781	21	AAV57453
33	115.5	5.5	2183	21	AAW03533
34	111.5	5.3	355	22	AAW40196
35	111.5	5.3	521	21	AAW57089
36	111.5	5.3	582	21	AAV44722
37	110.5	5.3	290	21	AAW43240
38	105	5.0	1070	22	AAW00894
39	100	4.8	830	21	AAV96730
40	99	4.7	558	22	AAW85479
41	99	4.7	2343	20	AAW80989
42	99	4.7	2343	21	AAV57846
43	98.5	4.7	700	12	AAW11354
44	98	4.7	804	21	AAW42128
45	97	4.6	340	21	AAW15554

ALIGNMENTS

RESULT 1
AAV67947
ID AAV67947 standard; Protein; 405 AA.
XX
AC AAV67947;
XX
DT 27-MAR-2000 (first entry)
XX
DE Human DETH protein SEQ ID NO:2.
XX
KW Human; DETH; apoptosis; TNFR; neurodegenerative diseases; cancer;
KW death domain expressing tumour necrosis factor receptor family homologue;
KW autoimmune disease; diagnosis.
XX
OS Homo sapiens.
XX
PN W09966039-A1.
XX
PD 23-DEC-1999.
XX
PF 08-JUN-1999; 99WO-CB01793.
XX
PR 12-JUN-1998; 98GB-0012607.
XX
(ZENEC) ZENEC LTD.
XX
Lu JJ, Gomes BC, Fieles WE;
WPI: 2000-097744/08.
N-PSDB; AA57184.
DR
PT New protein having 80% identity to death domain expressing tumor
PT necrosis factor receptor family homologue -
XX

Human protein sequ
Mouse TNF receptor
Human TNF receptor
Murine osteoproteg
A human tumour nec
Tumour necrosis fa
Human TNF receptor
Peptide #1918 enco
Peptide #3381 enco
Peptide #2009 enco
Peptide #3471 enco
Peptide #1923 enco
Peptide #3324 enco
Human TNF receptor
Mouse TNF receptor
Peptide #3583 enco
Peptide #3526 enco
Human TNF receptor
Human transcriptio
Murine factor V SE
Human polypeptide
Human prostate can
Human immune syste
Human ORFX ORF2004
Human bone marrow
PRO536, a Costal-2
Human serine/threo
Canine factor VIII
Canine Receptor VIII
Cellular Receptor
Human ORFX ORF1892
Apoptosis related

PS Claim 5; Page 35-36; 40pp; English.

CC The present sequence represents a human death domain expressing tumour
CC necrosis factor receptor family homologue, designated DEH. DEH has a
CC role in apoptosis and appears to be member of the TNFR1/FAS/NGFR family
CC of receptors. The DEH protein can be used for inducing apoptosis by
CC expressing DEH in a cell. The protein is useful for identifying
CC inhibitors and agonists of apoptosis that treat neurodegenerative
CC diseases (inhibition), cancer and autoimmune diseases (induction).
CC DEH-specific antibodies are useful for diagnosis of conditions and
CC diseases associated with DEH expression. The protein avoids the use
CC of harmful chemotherapy, where the loss of the p53 tumour repressor
CC gene can lead to drug resistant tumour cells following treatment.

XX Sequence 405 AA:

Query Match 100.0%; Score 2091; DB 21; Length 405;
Best Local Similarity 100.0%; Pred. No. 5,7e-197;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNTESNSASVPRKVLSTQEGVDPDNTSSARGKEPVNKTLPNLQVNHQGPVHRHIL 60
DB 1 mntesnsasvprkvlstqegvdpdntssargkevknktpnlqvyvnhqgpvhnhil 60
QY 61 KLPSMEATGEGKSTPIKPKRGHPRQNLKHFDTNEHLPMVILFLVLVIVVCS1 120
DB 61 klpsmeatggekstpikgprghprqnlkhfdtnehlpmviltlllvlvivvcs1 120
QY 121 RKSSRLTKKPRDPPSAIVKAGLKSMTPTQNRKRYTCNGHIDILKVAQVSSQW 180
DB 121 rkssrltkkprdpdpsaivkaglksmptqnrkrytcnghidiilkvaqvsqw 180
QY 121 RKSSRLTKKPRDPPSAIVKAGLKSMTPTQNRKRYTCNGHIDILKVAQVSSQW 180
DB 121 rkssrltkkprdpdpsaivkaglksmptqnrkrytcnghidiilkvaqvsqw 180
QY 181 KDLYOFLCNSEREVAFAFNGYTADEHRAVAALOHMTIRPEASLAOLISALROHRENDV 240
DB 181 kdlyoflcnserevaafngytagheraavaalohmtirpeaslaolisalrohren 240
QY 181 kdlyoflcnserevaafngytagheraavaalohmtirpeaslaolisalrohren 240
DB 181 kdlyoflcnserevaafngytagheraavaalohmtirpeaslaolisalrohren 240
QY 241 VEKIRGLMEDTTOLETDKIALPMSPPSPSPSPNAKLNSALLVEPSPQDKNGFF 300
DB 241 vekirglmedttoletdkialpmsppspspspnaaklensalllvepspqdkngff 300
QY 301 VDSESEPLLRCDNRSSGSSALSRRNGSFTTRKKTQVLRQVLDPCDDLPITDDMLHFNPE 360
DB 301 vdsesepilrcdnrsgssalsrrngsfttrkktqvlrqvldpcddlpitddmlhfnpe 360
QY 301 vdsesepilrcdnrsgssalsrrngsfttrkktqvlrqvldpcddlpitddmlhfnpe 360
DB 301 vdsesepilrcdnrsgssalsrrngsfttrkktqvlrqvldpcddlpitddmlhfnpe 360
QY 361 ELRVIEEIPQAEKLDLRFETIGVKSQEAQTLDSVYSHLPDL 405
DB 361 elrvieeipqaekldlrfetigvksqeaqtlldsvyshlpdl 405

RESULT 2
ID AAY05678 standard; protein; 631 AA.

AC AAY05678;

XX 19-JUL-1999 (first entry)

DE Human full-length tumour necrosis factor receptor ZTNFR-6.

KW ZTNFR-6; tumour necrosis factor receptor-6; human;
XX cell maturation; bone cell regulation.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..17 /note="signal peptide"

FT Protein 18..631 /note="mature protein"

FT Protein 18..188 /note="a polypeptide comprising this region of the
sequence is claimed in Claim 28(b)"

FT Query Match 100.0%; Score 2091; DB 20; Length 631;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;

FT Protein 18..325 /note="a polypeptide comprising this region of the
sequence is claimed in Claim 28(c)"

FT Protein 37..188 /note="a polypeptide comprising this region of the
sequence is claimed in Claim 28(a)"

FT Protein 66..145 /note="a polypeptide comprising this region of the
sequence is claimed in Claim 26(a)"

FT Region 37..65 /note="extracellular, cysteine-rich repeat"

FT Region 66..107 /note="extracellular, cysteine-rich repeat"

FT Region 108..145 /note="extracellular, cysteine-rich repeat"

FT Region 146..188 /note="extracellular, cysteine-rich repeat"

FT Region 189..325 /note="linker region"

FT Domain 326..346 /note="transmembrane domain"

FT Domain 347..631 /note="cytoplasmic domain"

FT Domain 390..471 /note="death domain, specifically claimed in
Claim 32(a)"

FT MO9911790-AI.

PD 11-MAR-1999.

XX 03-SEP-1998; 98WO-US18364.

XX 04-SEP-1997; 97US-0923725.

XX 04-SEP-1997; 97US-0057608.

XX (ZYMO) ZYMOGENETICS INC.

XX Farrah TM, Gross JA, Matthews SM;

PI WPI: 1999-205190/17.

DR N-PSDB; AAX25259.

XX New secreted or membrane bound tumor necrosis factor receptor

PT ZTNFR-6 - useful for detecting a genetic abnormality in a patient

PS Claim 26; Page 113-115; 145pp; English.

XX The present sequence represents a novel human full-length

CC membrane-bound tumour necrosis factor receptor, designated ZTNFR-6,

CC that is characterised by 4 cysteine-rich pseudo-repeat motifs.

CC ZTNFR-6 polypeptides can be obtained using recombinant techniques.

CC A polynucleotide (see AAX25259) encoding ZTNFR-6 has been isolated

CC from an EST database. A secreted, soluble form (see AAY05679) of

CC ZTNFR-6 has also been identified, which lacks the transmembrane

CC and cytoplasmic domains of the membrane-bound protein. ZTNFR-6

CC polypeptides, including the isolated extracellular region,

CC transmembrane domain and death domain, are claimed. ZTNFR-6

CC polypeptides are useful in methods that promote cellular maturation

CC and bone cell regulation. Antibodies raised against ZTNFR-6 are

CC useful for immunohistochemical tagging of cells expressing ZTNFR-6

CC for use in diagnosis, isolating ZTNFR-6, generating anti-idiotypic

CC antibodies, and as neutralising antibodies or antagonists that block

CC ZTNFR-6 in vivo and in vitro. The polypeptides can also be used to

CC identify agonists and antagonists of ZTNFR-6. Agonists are useful

CC for modifying the proliferation and development of target cells in

CC vitro and in vivo.

XX Sequence 631 AA;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MNSTESNSASVRPKVLSIQECTVPDNTSSARGKEDVKNKTLPNLOVNVHOGCPHRRHL 60
   |||
DB 227 mstesnsasvrpkvlsiqegtvpdntssargkedvknktpnlqvnhqgphrhhl 286
   |||
OY 61 KLLPSMEATGGEKSSPPKRGHPRONLHKHFDINEHLPMWIVFLVLVIVVCS1 120
   |||
DB 287 klpsmeatggekssppkrgphrpnllhkfhfdinehlpmwivflvlvlvvcsl 346
   |||
OY 121 RKSSRTLKKGPRDPSAIVEKAGLKKSMPTONREKWIYYCNGHGIDILKLVAAVGSQW 180
   |||
DB 347 rkssrtlkkgprdpasavekaglkksmptongrkwiiyycnghgidilkvvaqysqw 406
   |||
OY 181 KQIYQFLCNASEREVAFNSGYTADEHRAVALQHWITIGPPASLQOLISALROHRRNDV 240
   |||
DB 407 kqiyqflcnaserevaafnsygtadherayaalqhwitigppaslaqlisaltqrhrndv 466
   |||
OY 241 VEKIRGLMEDTTOLETDKLLALPMSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 300
   |||
DB 467 vekirglmedttqletdcklalpmsspsspsspsspsspsspsspsspsspsspdknkgff 526
   |||
OY 301 VDESEPLRCDSYSSGSSALSRRNGSFTTKEKKDVTLRQVRLDPCDLOPIFDDMLHFLNPE 360
   |||
DB 527 vdesepllrcdstsgssalsrrngsfttkekkdtvlrqrldpcdiqlpiddmlhflnpe 586
   |||
OY 361 ELRVIEIIPAEDKLDLRFETIGVKSOEASQTLDSVYSHLPDL 405
   |||
DB 587 elrvieeipgaedkldrlfeilgvsqeaasqtlldsvshlpdl 631
   |||
```

RESULT 3

AAW75792 standard; Protein: 655 AA.

```
XX AC AAW75792;
XX DT 21-DEC-1998 (first entry);
XX DE Human tumour necrosis factor related receptor TR7.
XX KM Tumour necrosis factor related receptor; TR7; human; inflammation;
XX KM arthritis; septicaemia; autoimmune disease; psoriasis; host
XX KM inflammatory bowel disease; infection; graft-versus-host disease;
XX KM transplant rejection; stroke; acute respiratory disease syndrome;
XX KM ischaemia; restenosis; brain injury; AIDS; bone disease; cancer;
XX KM atherosclerosis; Alzheimer's disease; diagnosis; therapy; vaccine.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 24 /note="encoded by ACG"
XX PN EP869179-A1.
XX PD 07-OCT-1998.
XX PF 01-APR-1998; 98BP-0302528.
XX PR 28-OCT-1997; 97US-0959382.
XX PR 02-APR-1997; 97US-0041796.
XX PA (SMK ) SMITHKLINE BEECHAM CORP.
XX PI Deen KC, Hurle MR, Tan KB, Young P;
XX DR WPI: 1998-508493/44.
XX DR N-PSDB: AAW57441.
XX PT New tumour necrosis factor receptor TR7 polypeptides and
XX PT polynucleotides - useful as diagnostic reagents and for treating
XX PT Alzheimer's disease, AIDS and cancer
```

XX Claim 11: Page 19-21; 25pp; English.

CC This is the amino acid sequence of a novel member of the human
CC tumour necrosis factor (TNF) receptor superfamily, termed TR7.
CC It was deduced from the coding regions of overlapping isolated
CC cDNA clones (see AAW57441). Expression systems, host cells and a
CC method of producing TR7 polypeptides are claimed. TR7 polypeptides
CC are useful for diagnosing diseases or susceptibility to diseases by
CC determining TR7 polypeptide or mRNA expression. TR7 polypeptide
CC can be used to screen for agonists and antagonists which bind the
CC receptor. These can be used in treatment to inhibit or enhance TR7
CC activity. TR7 antibodies are generated using TR7 polypeptide
CC fragments, and are used for treatment of diseases. TR7 polypeptides
CC and polynucleotides can be administered directly as vaccines for
CC prevention of diseases. Diseases diagnosed, treated or prevented
CC by the above methods include: chronic and acute inflammation,
CC arthritis, septicaemia, autoimmune diseases (e.g. inflammatory bowel
CC disease, psoriasis), transplant rejection, graft vs. host disease,
CC infection, stroke, ischaemia, acute respiratory disease syndrome,
CC restenosis, brain injury, AIDS, bone diseases, cancer (e.g.
CC lymphoproliferative disorders), atherosclerosis, and Alzheimer's
CC disease.

SQ Sequence 655 AA;

Query Match 100.0%; Score 2091; DB 19; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MNSTESNSASVRPKVLSIQECTVPDNTSSARGKEDVKNKTLPNLOVNVHOGCPHRRHL 60
   |||
DB 251 mstesnsasvrpkvlsiqegtvpdntssargkedvknktpnlqvnhqgphrhhl 310
   |||
OY 61 KLLPSMEATGGEKSSPPKRGHPRONLHKHFDINEHLPMWIVFLVLVIVVCS1 120
   |||
DB 311 klpsmeatggekssppkrgphrpnllhkfhfdinehlpmwivflvlvlvvcsl 370
   |||
OY 121 RKSSRTLKKGPRDPSAIVEKAGLKKSMPTONREKWIYYCNGHGIDILKLVAAVGSQW 180
   |||
DB 371 rkssrtlkkgprdpasavekaglkksmptongrkwiiyycnghgidilkvvaqysqw 430
   |||
OY 181 KQIYQFLCNASEREVAFNSGYTADEHRAVALQHWITIGPPASLQOLISALROHRRNDV 240
   |||
DB 431 kqiyqflcnaserevaafnsygtadherayaalqhwitigppaslaqlisaltqrhrndv 490
   |||
OY 241 VEKIRGLMEDTTOLETDKLLALPMSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 300
   |||
DB 491 vekirglmedttqletdcklalpmsspsspsspsspsspsspsspsspsspsspdknkgff 550
   |||
OY 301 VDESEPLRCDSYSSGSSALSRRNGSFTTKEKKDVTLRQVRLDPCDLOPIFDDMLHFLNPE 360
   |||
DB 551 vdesepllrcdstsgssalsrrngsfttkekkdtvlrqrldpcdiqlpiddmlhflnpe 610
   |||
OY 361 ELRVIEIIPAEDKLDLRFETIGVKSOEASQTLDSVYSHLPDL 405
   |||
DB 611 elrvieeipgaedkldrlfeilgvsqeaasqtlldsvshlpdl 655
   |||
```

RESULT 4

AAAY41693 standard; Protein: 655 AA.

```
XX AC AAY41693;
XX DT 07-DEC-1999 (first entry);
XX DE Human PRO868 protein sequence.
XX KM Human: PRO; EST; expressed sequence tag; PCR primer; hybridisation;
XX KM probe; blood coagulation disorder; cancer; cellular adhesion disorder;
XX KM secreted protein; transmembrane protein.
```


PR 24-MAR-1999; 9905-0126019.
PR 14-MAY-1999; 9905-0134220.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Gentz RL, Yu G, Fan P;
XX
XX WPI: 2000-594575/56.
DR N-PSDB: AAA99658.
XX
PT Nucleic acid molecule encoding a human tumor necrosis factor receptor,
PT known as TR9, useful for treating, preventing and diagnosing severe
PT combined immunodeficiency, autoimmune diseases, HIV infection, epilepsy
PT and cancer -
XX
XX Claim 20; Fig 1; 220pp; English.
XX
XX The present sequence is a novel human tumor necrosis factor
CC receptor, designated TR9. The TR9 receptor is also known as Death Domain
CC Containing Receptor 6. TR9 polypeptides, polynucleotides or agonists are
CC useful for treating, preventing or diagnosing common variable
CC immunodeficiency, X-linked agammaglobulinemia, severe combined
CC immunodeficiency and Wiskott-Aldrich syndrome, autoimmune diseases (such
CC as rheumatoid arthritis, allergic encephalomyelitis, multiple sclerosis,
CC diabetes mellitus and asplenia), HIV infection, epilepsy, cancer,
CC cardiovascular diseases and other neurological diseases.
XX
XX Sequence 655 AA:
SQ
Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTKLPMLQVNVHNOGPHRHIL 60
DB 251 mntesnsasvprkvlssiqegtvpdntssargkedvntklpmlqvnhqgphrhil 310
QY 61 KLPSPMEATGEGKSTPIKGRKRGHPRONLKHHPDINEHLPMTVLFLLVLVVWCSI 120
DB 311 klpsmeatggekstpikgrkrgpqnlnkhhdinehlpmtvlflvlvvcsl 370
QY 121 RKSRTLKKGRPDPSAIVEKAGLKKSWTPQNREKWTYYCNGHIDILKLVAAQVGSQW 180
DB 371 rksrtlkkgrpdpsaivekagllkswtpqnrkwtlycngnhydilkvvaqvgsw 430
QY 181 KDIYOFLCNMSERVEAFAFNSNGYTDHREAVYALOHWTIRGPEASLAOLISALROHRRNDV 240
DB 431 kdlyoflcnserevaafnsngytdhreyaalohwtirgpeaslaqlisallqrndv 490
QY 241 VEKIRGIMEDTTOLETDKLLPMSPPSLSPSPIPSPNAKLNSALLTVESPPODKNGFF 300
DB 491 vekirgimedttletdkllpmsppslspspipspnaksalltvepdpdkngff 550
QY 301 VDSEEPFLRCDSTSSGSSALSRRNSFTTKEKKOVLRQVRLDPDLDPIRFDMLHFLNPE 360
DB 551 vdseepflrcdstssgsalsrrnsfttkekkdvlrqlvrlpdcldpifdmlhflnpe 610
QY 361 ELRYIEEIPQAEKILDRLEFIIGVKSQEOASOTLLDSVYSHLPDLL 405
DB 611 elryieepqaedkildrlefiigvksqeoastlldsvyshlpdll 655
RESULT 7
ID AAB01349 standard; Protein; 655 AA.
XX
XX AAB01349:
XX
XX 20-OCT-2000 (first entry)
XX
XX Tumour necrosis factor receptor homologue TRH1 (Clone 2733717).
XX

KW Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW transplant rejection; activation; proliferation; differentiation;
KW apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW probe; primer; human.
XX
OS Homo sapiens.
XX
XX WO200034294-A2.
XX
XX 15-JUN-2000.
XX
XX
XX 10-DEC-1999; 99WO-US29400.
XX
XX 11-DEC-1998; 98US-0111826.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Bowen MA, Siemers N;
XX
XX WPI: 2000-423364/36.
DR N-PSDB: AAA47395.
XX
XX Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT development
XX
XX Claim 10; Fig 2a-2d; 42pp; English.
PS
XX The tumour necrosis factor receptor homologue TRH1 can be used for
CC treating a mammal e.g. a human, at risk for a disorder characterized
CC by an aberrant or unwanted level or biological activity of TRH1,
CC e.g. Rheumatoid arthritis and transplant rejection. TRH1 may also be
CC useful to leach out or block a ligand which is found to bind to the
CC TRH1. TRH1 may be used in various drug screening techniques and to
CC identify fragments and analogs of a protein or peptide (agonist or
CC antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC cellular function, cell activation, proliferation, differentiation,
CC and apoptosis. The interaction between the novel TNFR protein of the
CC present invention and intracellular signaling molecules and/or its
CC potential co-receptor may serve as a novel target for
CC immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC development. Gene constructs can also be used as part of a gene
CC therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC directed against TRH1 can be used to reject TRH1 in tissues
CC and cells. They can also be used to make targeted antibody that
CC destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC used as diagnostic probes or as PCR primers. Fragments of the full
CC length gene may be used as hybridization probes for a cDNA library to
CC isolate the full length gene and to isolate other genes which have a
CC high sequence similarity. The probes may be used to identify a cDNA
CC clone corresponding to a full length transcript and a genomic clone
CC or clones that contain the complete gene including regulatory and
CC promoter regions, exons, and introns.
XX
XX Sequence 655 AA:
SQ
Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNSTESNSASVPRKVLSSIOEGTVPDNTSSARGKEDVNTKLPMLQVNVHNOGPHRHIL 60
DB 251 mntesnsasvprkvlssiqegtvpdntssargkedvntklpmlqvnhqgphrhil 310
QY 61 KLPSPMEATGEGKSTPIKGRKRGHPRONLKHHPDINEHLPMTVLFLLVLVVWCSI 120
DB 311 klpsmeatggekstpikgrkrgpqnlnkhhdinehlpmtvlflvlvvcsl 370
QY 121 RKSRTLKKGRPDPSAIVEKAGLKKSWTPQNREKWTYYCNGHIDILKLVAAQVGSQW 180
DB 371 rksrtlkkgrpdpsaivekagllkswtpqnrkwtlycngnhydilkvvaqvgsw 430

XX The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HDTE84, HSLJD37R and RANKL;
CC human CC chemokine HCC5; human deubiquitinating proteins Dab1 and Dab
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif
CC (LRR); human cyclin E2; CDNA encoding these proteins; and antibodies
CC against these proteins. The proteins can be used for modulating the
CC physiology or development of a cell. They can be used to mediate uptake
CC of substrates (e.g., prostaglandin-like molecules), to modulate or
CC mediate cellular interactions (e.g., induce or prevent trafficking,
CC proliferation, or differentiation of cells), or are intracellular
CC proteins which are important in various cellular processes such as the
CC deubiquitination of proteins or cell cycle regulation. The products can
CC be used for treating medical conditions such as immune, inflammatory or
CC allergic disorders, or abnormal cellular proliferation, for example,
CC cancers or degenerative conditions. They can be used to modulate immune
CC responses in disease states e.g., autoimmune disorders, including
CC rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC autoimmune thyroiditis, as well as acute and chronic inflammatory
CC responses in which T cell activation, expansion, and/or immunological T
CC cell memory play an important role. Sequences AAY77458-177461 and
CC AAY77465-177468 represent TNF receptor family-related proteins. AAY77458
CC is the human protein HDTE84, AAY77459-177461 are human HSLJD37R
CC proteins, AAY77465 is murine Rank-like protein RANKL, and AAY77466-177468
CC are human RANKL proteins.
XX
XX Sequence 655 AA:

Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSTESNSASAVRKVLSSIOEGVPPDNTSSARKEDEVNKLPLNQVNHQGGPHRHIL 60
DB 251 mstesnsasavrkvlssioegvppdntssarqedvnlplnqvnbgqphrhil 310
QY 61 KLPSMEATGEGKSTPKPKRGHPRONLHKHPDINHLPMMIYFLFLVLYVVCSEI 120
DB 311 klpsmeatgksscpkpkrgprgnlnkhhdinehlpmmiylfllvlyvvcse 370
QY 121 RKSSRTLKKGRPDPSAIVERAGLKKSMPTQNREKWIYVNCNGHIDILKLVAAQVGSQW 180
DB 371 rksrtlkkgrpdpsaivekagllksmtcqnrekwiycngidilklvaavgsw 430
QY 181 KDIYOFICNASEREVAANSNGYTADHERAVYALQHWITRGPEASTLAQLISLRQHRNDV 240
DB 431 kdiyoflcnaserevaansngytadherayaalqhwitrgpeastlaqlislrqhrndv 490
QY 241 VEKIRGLMEDTTOLETDKLALPMSPLSPSPPIPSNAKLNSALTYEPSPQKNKCF 300
DB 491 vekirglmedtltqetdclalpmespplspspipspnaklensalltyepspqknkcf 550
QY 301 VDESEPLLRCDSSTSGSSALSRLNGSFIRKEKKDVYLRQVRLDPCDLPIDFDMHLFLNPE 360
DB 551 vdesepllrcdstsgssalsrlngsfirkekkdvylrqvrlldpcdlpifdmlhflnpe 610
QY 361 ELRYIEERIPAQEDKIDRFELIIGVKSQASOTLDSVSHLPDL 405
DB 611 elryieeripaedkidrfeiliigvksqasotlldsvshlpdl 655

RESULT 10
AA92846
ID AA92846 standard; Protein; 691 AA.
XX
AC AA92846;
XX
DT 29-AUG-2000 (first entry)
XX

DE Human osteoprotegerin-like 4 (OPGx4) protein.
XX
XX OPGx4; osteoprotegerin-like; DR6 TNF-related death receptor; agonist;
KW inhibitor; bone resorption; vascular calcification; apoptosis;
KW osteopathic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..36
FT /note- "36 amino acid extension relative to
FT DR6 TNF-related death receptor"
XX
XX WO200024771-A2.
XX
XX 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US24913.
XX
XX 23-OCT-1998; 98US-0105481.
XX 01-OCT-1999; 99US-0156993.
XX 21-OCT-1999; 99US-0422680.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Yang M, Lichenstein H, McDonald WF;
XX
XX WPI: 2000-350692/30.
XX DR N-PSDB: AAA28728.
XX
XX Nucleic acids encoding osteoprotegerin-like (OPGx) polypeptides, useful
PT for treating disorders associated with bone metabolism, such as
PT osteoporosis and osteopetrosis
XX
XX Claim 24: Fig 2; 81pp; English.
XX
XX Osteoprotegerin-like polypeptides (OPGx) are expressed in bone, lymph
CC node, germinal B cells and kidney. There appear to be at least two splice
CC variants of this gene, a transmembrane form and an extracellular domain
CC form. Human OPGx1, 2 and 4 contain a 36 amino acid extension at their
CC N-terminal relative to the DR6 tumour necrosis factor (TNF)-related death
CC receptor. The OPGx polypeptides, agonists and antibodies are useful in
CC methods to inhibit osteoclast-mediated bone resorption or vascular
CC calcification and to modulate cell death (apoptosis). This is useful for
CC treating disorders associated with bone metabolism, such as osteoporosis,
CC osteopetrosis, or a condition characterized by loss of bone, breakdown of
CC tissue, or excessive readorption of bone tissue.
XX
XX
XX Sequence 691 AA:

Query Match 100.0%; Score 2091; DB 21; Length 691;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSTESNSASAVRKVLSSIOEGVPPDNTSSARKEDEVNKLPLNQVNHQGGPHRHIL 60
DB 287 mstesnsasavrkvlssioegvppdntssarqedvnlplnqvnbgqphrhil 346
QY 61 KLPSMEATGEGKSTPKPKRGHPRONLHKHPDINHLPMMIYFLFLVLYVVCSEI 120
DB 347 klpsmeatgksscpkpkrgprgnlnkhhdinehlpmmiylfllvlyvvcse 406
QY 121 RKSSRTLKKGRPDPSAIVERAGLKKSMPTQNREKWIYVNCNGHIDILKLVAAQVGSQW 180
DB 407 rksrtlkkgrpdpsaivekagllksmtcqnrekwiycngidilklvaavgsw 466
QY 181 KDIYOFICNASEREVAANSNGYTADHERAVYALQHWITRGPEASTLAQLISLRQHRNDV 240
DB 467 kdiyoflcnaserevaansngytadherayaalqhwitrgpeastlaqlislrqhrndv 526
QY 241 VEKIRGLMEDTTOLETDKLALPMSPLSPSPPIPSNAKLNSALTYEPSPQKNKCF 300
DB 526 vekirglmedtltqetdclalpmespplspspipspnaklensalltyepspqknkcf 586

```
DB 527 vekirgmedtcqlecdkialpmspslpspiipnakiensalltvepsqdknqgff 586
OY 301 VDESEPLLRCDSTSSSSALSRNGSFITTEKKDFTVLRQRLDPCDQPIFDMLHFLNPE 360
DB 587 vdesepllrcdstssgsalsrnsqflltkkdkvtlrvlrdpcdipifddmlhflnpe 646
OY 361 ELRVIEEIPQAEKLDRLFEITGVKSQEAQOTLLDSVYSHLPDL 405
DB 647 elrvieelpqaedkldrlfeitgvksqeaqotlldsvyshlpdl 691

RESULT 11
AAB44249
ID AAB44249 standard; Protein: 655 AA.
AC AAB44249;
DT 08-FEB-2001 (first entry)
DE Human PRO668 (UNQ437) protein sequence SEQ ID NO:64.
XX Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
KM expressed sequence tag; detection; cancer.
XX Homo sapiens.
PN WO200053756-A2.
PD 14-SEP-2000.
XX 18-FEB-2000; 2000MO-US04341.
PF 08-MAR-1999; 99MO-US05028.
XX 12-MAR-1999; 99US-0123937.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130233.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99MO-US28551.
PR 02-DEC-1999; 99MO-US28565.
PR 16-DEC-1999; 99MO-US30095.
PR 30-DEC-1999; 99MO-US31243.
PR 30-DEC-1999; 99MO-US31274.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00277.
PR 06-JAN-2000; 2000MO-US00376.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
PI Ferrara N, Fllvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI.
XX
DR WPI; 2000-611443/58.
DR N-PSDB; AAC78474.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities.
XX
XX
XX Claim 12; Fig 26; 636pp; English.
XX
XX AAC78458 to AAC76599 represent polynucleotide and EST (expressed
XX sequence tag) sequences which encode secreted or transmembrane PRO
XX polypeptides. The PRO polynucleotides and polypeptides have cytostatic
XX activity. The polynucleotides and polypeptides can be used for detecting
XX the presence of PRO polypeptides in samples, for linking bioactive
```

```
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 655 AA:

Query Match 99.9%; Score 2089; DB 21; Length 655;
Best local Similarity 99.8%; Pred. No. 1; 9e-196;
Matches 404; Conservative 1; Mismatches 0; Indels 0; Gaps 0:

OY 1 MNSTESNSASVPRKYLSTIQEOTVPDNTSSARGKEDVNTLRLNLOVYNNIQGPHRHIL 60
DB 251 mstesnsasvprkylstiqegvtvpdntssargkedvntlrlnlovyvnniqgphrhil 310
OY 61 KLLPSMEATGCEKSPPIKGRGHPRONLHKHFDINEHLPMWIVLFLVLVYVCS1 120
DB 311 klpsmeatgceksppikgrgprnlnkhfdinehlpmwivlflvlvlyvcs1 370
OY 121 RKSSRTLKKGRPDPSAIVEKAGLKKSMPTQNRKWIYVYCNHGIDILKLVAAVGSQW 180
DB 371 rkssrtlkkgrpdpsaivekaglkksmptqnrekwiycnbgidilklvaavgsw 430
OY 181 KDLYQELCNASEREVAFAFSNGYTADEHRAVALQHWITRPEASLAQJISALQOHRNDV 240
DB 431 kdlyqelcnaserevaafsfngyadheravalaqhwitrpeaslaqljlsalqhrndv 490
OY 241 VEKIRGIMEDTQLEMDKIALPMSPSPLSPSPISPSPAKIENSALLTVEPSPODKNKGFF 300
DB 491 vekirgimedtcqlecdkialpmspslpspiipnakiensalltvepspqdknqgff 550
OY 301 VDESEPLLRCDSTSSGSALSRLNGSFITTEKKDFTVLRQRLDPCDQPIFDMLHFLNPE 360
DB 551 vdesepllrcdstssgsalsrnsqflltkkdkvtlrvlrdpcdipifddmlhflnpe 610
OY 361 ELRVIEEIPQAEKLDRLFEITGVKSQEAQOTLLDSVYSHLPDL 405
DB 611 elrvieelpqaedkldrlfeitgvksqeaqotlldsvyshlpdl 655

RESULT 12
AAB93023
ID AAB93023 standard; Protein: 655 AA.
AC AAB93023;
DT 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:11787.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 17-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:37 ; Search time 15.98 Seconds

(without alignments)
1930.582 Million cell updates/sec

Title: US-09-733-956-2

Perfect score: 2091

Sequence: 1 MNSTRENSASVRPKVLST.....SQEASQFLDLSVSHLPDLL 405

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115.5	5.5	2183	2 T42764	coagulation factor
2	115	5.5	927	2 T51536	hypothetical prote
3	114.5	5.5	290	2 T08692	hypothetical prote
4	112	5.4	1822	2 T14106	probable GTPase-ac
5	107	5.1	309	2 T17557	procyclin homolog
6	107	5.1	461	2 T22946	hypothetical prote
7	105.5	5.0	452	2 T46147	zinc finger protel
8	105.5	5.0	524	2 E71881	hypothetical prote
9	105.5	5.0	1405	2 T04426	hypothetical prote
10	104.5	5.0	607	2 T40906	probable ATP-depen
11	104.5	5.0	3140	2 S47508	genome polyprotein
12	103.5	4.9	1301	2 S51323	SAC3 protein - yea
13	103	4.9	1231	2 T18532	serine/threonine pr
14	103	4.9	2176	2 T13806	lucan gene protei
15	101.5	4.9	494	2 T28067	hypothetical prote
16	101.5	4.9	514	2 T15338	hypothetical prote
17	101.5	4.9	527	2 B64633	hypothetical prote
18	101	4.8	342	2 E82955	tona protein PA553
19	100	4.8	498	2 P83523	probable colicin-1
20	100	4.8	1087	2 T49496	hypothetical prote
21	99.5	4.8	872	2 T50369	probable serine/th
22	99	4.7	604	2 T41249	DEAD box ATP-depen
23	98.5	4.7	3140	1 GNV5RA	genome polyprotein
24	98.5	4.7	26926	1 T38344	titin, cardiac mus
25	98	4.7	670	2 T13739	probable hormone r
26	98	4.7	887	2 S61137	probable membrane
27	98	4.7	1188	2 T20333	hypothetical prote
28	97	4.6	764	2 T05409	hypothetical prote
29	96.5	4.6	1643	2 T07961	myosin heavy chain

30	96.5	4.6	1940	2 T42715	ankyrin 3, splice
31	96.5	4.6	1943	2 T42713	ankyrin 3, splice
32	96.5	4.6	1941	2 T42716	ankyrin 3, splice
33	96	4.6	607	2 S27776	80K protein (allel
34	96	4.6	856	2 G71133	probable alpha-man
35	96	4.6	969	2 T41707	probable pheromone
36	95.5	4.6	733	2 S61984	probable membrane
37	95.5	4.6	962	2 A84497	probable retroelem
38	95.5	4.6	1025	2 S54044	probable membrane
39	95.5	4.6	1086	2 S16752	major merizolite su
40	95.5	4.6	1502	1 R6BYH1	CYC1/CYP3 transcri
41	95	4.5	707	2 S57157	hypothetical prote
42	95	4.5	3141	1 GNV5PD	genome polyprotein
43	94.5	4.5	1033	2 S19247	cell adhesion prot
44	94	4.5	249	2 T07300	hypothetical prote
45	94	4.5	294	2 G84504	probable VSP-1-lik

ALIGNMENTS

RESULT 1
T42764
coagulation factor V - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C:Accession: T42764
R:Yang, T.L.; Cui, J.; Rehmtulla, A.; Yang, A.; Mousaelli, M.; Kaufman, R.J.; Ginsbu
Blood 91, 4593-4599, 1998
A:Title: The structure and function of murine factor V and its inactivation by protei
A:Reference number: 222270; PMID:98282202
A:Accession: T42764
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2183 <YAN>
A:Cross-references: EMBL:052925; NID:93219690; PID:93219691; PIDN:AAC9953.1
C:Function:
A:Pathway: blood coagulation
C:Superfamily: coagulation factor V; disocidin I amino-terminal homology; ferroxidase
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm
F:350-682/Domain: ferroxidase repeat homology <FOX1>
F:1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match	Best Local Similarity	Score	Pred. No.	DB 2;	Length
Matches 97;	Conservative 49;	Mismatches 197;	Indels 145;	Gaps 20;	
5	ESNSASVRPKVLSSIQGVDP--NTSARGKEDVNTLPNL-----QVNHQO 52				
773	DSNS-----RLSKITNNLNKDQRLPGSGATVAGTLLRLNGDENFVLNSTEHR 827				
53	GPHRRH-----ILKLPSMEATGGEKSTPIKGRGPRONLHKHF-----D 95				
828	SSYHENDMENPQSNITWYLLPLPGKSGNREODKPTIKGRPHMKHRRSMKAPACK 887				
96	INELPMIIVFLLLVLVIVVCSIRKSRLLKGRPDPSAIEYKAGLKKSMPTQORE 155				
888	TGRH-----SNPKNSYSGMSEEDIPSELIP--LKOKITSKFLNR 925				
156	KW-----IYCCNGHGIDILKLVAAQGSQWKDIYFLCNASEREVAFAFNGYAD 205				
926	RWRVASEKSGYEITTAANGEDVDVRLT-----NSPONNTTVRGESTS 969				
206	HERAVAAALQHTWIRGPE---ASLAOLISALRHRNDVVEKIRGLMEDTTOLETKLA 260				
970	HTN-----TTRKPSDLPTFSGVGHKSPHVRDEENSGFQKROLFIRTRKKKKKKKLA 1021				
261	L--PMSPSPISP-----SPTEPNAKLKNSALL-----TYEPSPDKNKKFPVDESEP- 306				
1022	LHSPLSRPGFPLRGHNHNSPPDRRLNLHSLLRKSNETALSPDLNNTSPSMSTRSLPD 1081				
307	--LIRCDSTSSGSSALSRNGSFITREKKDQVLRQ-----VRLDCDQPIPD- 351				

Db 1082 YNOYSKNDEOMSSSLDYOSVAPAEHSPTPAQDPDQTHSTDPSSYSSPELSQGLDY 1141
 QY 352 DMLHNEPEELRVIEIEIPQAEKDL-----RLEIIGVSGQESQFLDLSY 397
 Db 1142 DLSHDFYDDGLTSFFPDQSKSSFSDDQAIPLSDLSFTI-----SPELDQTI---I 1194
 QY 398 YSHLPDL 405
 Db 1195 YPDDQL 1202

RESULT 2
 T51536
 hypothetical protein T20K14_190 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C:Accession: J51536
 R:Sato, S.; Nakamura, Y.; Kaneo, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
 submitted to the Protein Sequence Database, August 2000
 A:Reference number: 225394
 A:Accession: J51536
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-927 <SAR>
 A:Cross-references: EMBL:AL391143
 A:Experimental source: cultivar Columbia; BAC clone T20K14
 A:Genetics:
 A:Map position: 5
 A:Introns: 51/1; 69/3; 620/3; 668/1
 A:Note: T20K14_190

Query Match 5.5%; Score 115; DB 2; Length 927;
 Best Local Similarity 18.9%; Pred. No. 0.99;
 Matches 83; Conservative 51; Mismatches 158; Indels 146; Gaps 15;

QY 17 LSSIOEGTVPDN--TSSARGKEDVNTLPNLQVNHQGPNNHRLK----- 62
 Db 45 LKSLPSGKASDNDVGDNTNISADKKETKSKKKTKAKEQGVSESSRLSFSSPCSSSF 104
 QY 63 -----LPSMEATGGEKSSPTPIKGRGHR-ONLKHHPINELPMIYLFLLVLY 113
 Db 105 SSADISTTASQFEPDGLSNGENPVREPTNGSPRWGLMPSDIRE----- 149
 QY 114 VIVCSIRKSRKTL-KGPRDPSAIVEKAGLKSMPTQNEKKIYYCNGHIDILKLY 172
 Db 150 -LVRSIHKETRTREELLSQPKSARANVSLKSSPSRNSNM-----SGGRVYVKL- 202
 QY 173 AAQVSGQMKDIYQFLCNASEREVAAFNSNGYTDADHERAYALLOHWTIRGPEASLAOLISAL 232
 Db 203 -----KDSFRF--SYDERE----- 214
 QY 233 ROHRNDVVEKIRGLMEPTOLETDKIALPMSPLSPSPIP-----SPNAKL 280
 Db 215 -----TRTKGAFLKETPRRLSLDSRSNFSARSSCSPREQELVTGHRRTSSVYAKL 266
 QY 281 ENSALLTVEPSPDKNKGFVEDESEPLLRCDSTSSGSAISFNGSFTTEKK----- 332
 Db 267 MGLEVTIPPEPVYIQNRNENFCDSPPR-----TSRVEVDLQKSRGSDSKKKMKAPKFPK 320
 QY 333 -----DTVLQVRLDPCDLOPIFDDM-----LHFLNPF-ELRVIEIIPQAEKDLRL 378
 Db 321 ASPMAQVQKAKQKQVIRIPATTLTYVGEIQRLSQLEFKKSEKDLALAKIILEAMERTQOL 380
 QY 379 FEIIGVKQGEASQTLDS 396
 Db 381 IS-----KDDDKNTLCSS 394

RESULT 3
 T08692
 hypothetical protein DKFZp564k112.1 - human (fragment)

C:Species: Homo sapiens (man)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
 C:Accession: T08692
 R:Diesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, May 1999
 A:Reference number: 216470
 A:Accession: T08692
 A:Molecule type: mRNA
 A:Residues: 1-290 <DUE>
 A:Cross-references: EMBL:AL049996
 A:Experimental source: fetal brain; clone DKFZp564k112
 A:Genetics:
 A:Note: DKFZp564k112.1

Query Match 5.5%; Score 114.5; DB 2; Length 290;
 Best Local Similarity 21.1%; Pred. No. 0.22;
 Matches 66; Conservative 46; Mismatches 104; Indels 97; Gaps 13;

QY 2 NSTESNSASVAPKVLSSIOEGTVPDNTSSARGKEDVNTLPNLQVNHQGPNNHRLK 61
 Db 6 SSSEDAESLAPRSKYKLIQKDLIMEASEASKENDI-----ELSPTEE----- 49
 QY 62 LLPMEATGGEKSSPTPIKGRGHRONLKHHPDINEHP-----WAIYLFLLVLY 113
 Db 50 -----KDTGDLKDSLLK-TKRKHKKHKEHKGEEVPLPLVLSKSEWMDLKEYLALQ 103
 QY 114 VIVCSIRKSRRLK-KGPRDPSAIVEKAGLKSMPTQNEKKIYYC-----NGHGI 166
 Db 104 KASMASLKTITQIKSESMEITDGVPTGKKN--KTANEE---CRTEKKNATGP 157
 QY 167 D-----TLKVAQVSGQMKDIYQFLCNASE-----REVAAFSNG 201
 Db 158 QPVSGIYVKIISTEPLPGRKQVRDLAISEVLYDLLEGTECHARFKTPDAQAVINA 217
 QY 202 YI-----ADHERAYALLOHWTIRGPEASLAOLISALROHRNDVEKIRGL 247
 Db 218 YTEINKRHQCKLEILSGDEORY---W-----QKILVDROAKLNOPREKRGOT 262
 QY 248 MEDTQOLETDKLA 260
 Db 263 EKLTIOAEKIRLA 275

RESULT 4
 T14106
 probable GTPase-activating protein SPA-1 - rat
 N:Alternate names: protein p1294
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14106
 R:Takeuchi, M.; Ide, N.; Hata, Y.; Takai, Y.
 submitted to the EMBL Data Library, September 1997
 A:Description: SPA-1 like protein identified through yeast two-hybrid screening using
 A:Reference number: Z17877
 A:Accession: T14106
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1822 <TAK>
 A:Cross-references: EMBL:AF026504; NID:g2555182; PID:g2555183; PIDN:AA881526.1

Query Match 5.4%; Score 112; DB 2; Length 1822;
 Best Local Similarity 20.0%; Pred. No. 4.3;
 Matches 91; Conservative 65; Mismatches 152; Indels 148; Gaps 19;

QY 2 NSTESNSASVAPKVLSSIOEGTVPDNTSSARG--KEDVNTLPNLQVNHQGPNNHRI 59
 Db 1429 NSTFSINDATSHSTWSS-RHSASPVVFSARSSPREELHPTSS----- 1472
 QY 60 LKLLPMEATGGEKSSPTPIKGRGHRONLKHHP----- 94
 Db 1473 -GLAPFS--SSSSSGPRTEYPRQAGATSKIGKKRPGDTINSVGFMTRKHOS 1526


```

QY 121 RKSSTLTKKG-----PRDDP-AIVERAKGLKSKSTPTPONREKKIYYCGHGIDILK 170
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
Db 93 QKSNEKVKKKYYCVCEVSCVHHDPRALGDLTGKHKHCKRKHGEKKW--KCD---KCSK 146
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
QY 171 LVAAGVSGMKDIYOFLOCNASEREV--VAFSNGYTAHERAYA-ALOHWTIRGEASIA 226
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
Db 147 KYAAV--SDMK-AHSKIGCTKEYKCDGTLFSRDSFTTHRAFCALAEENARSHS--- 200
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
QY 227 QLISLRQH-----RRNVYVEKIRLMEDTQLELDKIALPRSPGLSPSPINAKLE 281
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
Db 201 ---QSKKNPELTRKNPV-----PNPV-PAPVDTESAKIK 232
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
QY 282 NSALLTVESPQD-----NSALLTVESPQD-----NKG 298
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
Db 233 SSSFTLRKQSESPKPPPIVQEPKPTSLUNVYTSNGVFAGLEPSSASPISYTTSSSKS 292
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
QY 299 FFVDES--EPLLRCDSTSSGSALS RN 323
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|
Db 293 LFASSSTIEPISLGLSTHSGSEFLGN 319
   :|::: :|| 11111 : 11111 :|::: :|| 11111 :|

```

RESULT 8

hypothetical protein jhp0842 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: E71861
R:Altun, R.A.; Lung, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 1997, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
A:Reference number: A71800; MUID:99120557
A:Accession: E71861
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-524 <ARN>
A:Cross-references: GB:AE001513; GB:AE001439; NID:g41155407; PIDN:AA006420.1; PID:g4155407
A:Experimental source: strain J99
C:Genetics:
C:Gene: jhp0842

Query Match	5.08;	Score 105.5;	DB 2;	Length 524;
Best Local Similarity	18.88;	Pred. No. 2.4;		
Matches 93;	Conservative 67;	Mismatches 180;	Indels 155;	Gaps 19

```

QY 2 NSTESNSASVAPKVLSSIOECTVDPNTSSAGKCDVKNCTPLNQLVNHQOCPHRRHLK 61
   1 : | | : : : : : : : : | : | : |
Db 11 NANALNSAKKNEVDKAKNAKPSASKDFSKILNOKISTKDTAP-----KESPNI-SALK 62
QY 62 LLP---SMEATGGEKST-----PIKPKRGHPRONLJHKHF 94
   1 : | | : : : : : : : : | : | : |
Db 63 ATRPKADKADAKLEKTRPLPHQHAONLAKDQAPTRKLDWLNRPKNHPRAPHAQ--HEHN 120
QY 95 DINEHLPMMIVLFLLLVAVVYVCSIRKSSRTLKGGPRODSAIYEKA-----142
   1 : | | : : : : : : : : | : | : |
Db 121 EANENTP-----KTPNETLSKNEKK-PNEVASNAHQTNLPKNKPNIT 160
QY 143 -----GLKKSMTPTQNBREKWTYCNCGHIDILKLVAAQVGSQWKDIQYFLCNASEREVA 196
   1 : | | | | : : : : : : : : | : | : |
Db 161 PNHANNAKIKNPTPHNVK-----DKTKLKDIQYLSQKHDL-----NASNTQAA 204
QY 197 AFSNGYTAADHERAVYALQHWITRGPEAST-----AQLISAL-----RO 234
   1 : | : : : : : : : : | : | : |
Db 205 T-----TEENKNPLASDHLALKTQASINHTLTNDKANKTANLSSVLOSLEKKEPÖNKE 259
QY 235 HRRNDVVERIKRGLMEDTQOLEFDKLALPMSPSPSPSP-----SPNAKLENS-----283
   1 : | : : : : : : : : | : | : |
Db 260 HANPNNEKKTPLPEALQÖMNAIKRDKTLSSKKKPKPKPTPIHAKTQTLIASVPENNAKPIPL 319
QY 284 -----ALLVEPSPÖD-----KNKGFEVDESEPLLRCDSITSSGSSALSIRNGSFTIK 329

```

Db 320 KTLPLMPLIGANPPNDNAPTLEKEETKEISDNKKK--ETNSSAQSVNQNTQASDKTS 378

QY 330 EKKDLY-----LEQVRIDPCDLOPIFDDMLHFLNPEELRYTEELPQAEKDKIDRLFEI 381

Db 379 DNKSTPKETIRHFTQOLKOEIOEYKPPMSKISMDLEPKELGVEYTI--KVGKNLKV 435

QY 382 IGVQSGEASOTLIDS 396

Db 436 SVISHNNSLQTELDN 450

RESULT

hypotheetical protein T18B16.20 - Arabidopsis thaliana
 N:Alternate names: hypotheetical protein F13C5.220
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
 C:Accession: T04426; T05042
 R:ByRef, M.; Banes, V.; Reichmann, S.; Borkove, D.; Ansgorge, W.; Bancroft, I.; Mewes, submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15359
 A:Accession: T04426
 A:Molecule type: DNA
 A:Residues: 1-1405 <BEV>
 A:Cross-references: EMBL:AL021687
 A:Experimental source: Cultivar Columbia; BAC clone T18B16
 R:ByRef, M.; Pohl, T.; Weizengeger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schmitt, submitted to the Protein Sequence Database, February 1998
 A:Reference number: Z15395
 A:Accession: T05042
 A:Molecule type: DNA
 A:Residues: 1-1405 <BEW>
 A:Cross-references: EMBL:AL021711
 A:Experimental source: Cultivar Columbia; BAC clone F13C5
 C:Genetics:
 A:Map position: 4
 A:Introns: 893/3; 1164/2; 1183/3; 1192/3; 1198/2; 1214/3; 1251/1; 1282/2; 1327/1; 1381/1
 A:Note: T18B16.20; F13C5.220

Query Match	5.0%;	Score 105.5;	DB 2;	Length 1405;
Best Local Similarity	22.1%;	Pred. No. 9.5;		
Matches 93;	Conservative 69;	Mismatches 156;	Indels 103;	Gaps 21;

```

QY 17 LSSIQGVPYDQNTSSARGKEDY-----NKPLNJOVYVNHQGGPHNRHILKLSMSEMT 69
Db 756 ISETNLSELPOKXISELSNLIKELIRKCSKLTPLPWDEKITN-----LEIFDVS 803
QY 70 GGEKSTPIKGGKRGHPRONLHKHFDINE---HLPWMLVLFLLVAVYVCSIRKSSR 125
Db 804 GCTELET-IEGSFEN--LSCLHK-VNLSETNLGELPNKXISELSNLIKELIRKCSKLTALP 859
QY 126 TLKKGPRDDPSAIYKAGLAKSMPTQONREKITYIYCNHGIDILKLVAAQV----SOWK 181
Db 860 NLEK---LTHVLIFVSGCTNLDKIEESEFESMSYLCESIYSSKRIYADSSCIEDOMS 916
QY 182 DLYOFLCANSEKVEVAFSSGYADHERAVAALOQHTITGCPKASLQOLLSALHQHRRQV 241
Db 917 QIKE--CLTSKKEGSSFSNVGKETREKLLYHGNRRYVDPEVPL-----NIDIV 963
QY 242 EKIRGLMEDTTOLEFDKDALPMSPPSLSPSPRIPEENAKLENSALTVEPSPODKRKGFEV 301
Db 964 D-----IKRSTDLKIEYIAKAEYVS-IAENGSKSVSLFEDLOQMSV-----KGMV 1009
QY 302 D-----ESEPLLRCDSITSSGS-----SALSRRGSPITKEKKDTVLVR 339
Db 1010 ERCKNMADVLEESDEDOLEKREKSSPSLOTLMTLSNLPILTSYSSKGGFIFKN-----LKKL 1064
QY 340 RUDPC-DLOPTDDMLHPLNPELAVIEETPOAEKDLORLFFIIVKSGQEAQT-----LL 394
Db 1065 SYVDCDPSIKWLPPEL-----PDLNLELR--VFQCKLLELFE--VKAGELSKRLKHL 1114

```

QY 395 D 395
Db 1115 D 1115

RESULT 10

T40906
probable ATP-dependent DNA helicase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T40906
R:Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, December 1998
A:Reference number: 221956
A:Accession: T40906
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-607 (MUR)
A:Cross-references: EMBL:AL034490; PIDN:CAA22471.1; GSPDB:GN00068; SPDB:SPCC126.02c
A:Experimental source: strain 972h; cosmid c126
C:Genetics:
A:Gene: SPDB:SPCC126.02c
A:Map position: 3
A:Introns: 11/3; 72/3; 105/1; 501/3; 571/3
C:Superfamily: thymid autoantigen

Query Match 5.08; Score 104.5; DB 2: Length 607;
Best Local Similarity 20.18; Pred. No. 3.5; Mismatches 158; Indels 131; Gaps 24;
Matches 91; Conservative 73; Mismatches 158; Indels 131; Gaps 24;

QY 34 GKEDVAKTLPNLOVNHQGPRIHRIKLPSEATGCEKSPYIKGPKGHPRONLHKH 93
Db 204 GRDVS-----NLVNRGGA-QLOHMLMT-----TALQPKRAH-----FLK 240
QY 94 FDIENHLPWMIYVLELLVNVIVVCSIRKSRLLKGRP--ODPSAIVKAGLKSMTP 150
Db 241 MDGNOVRIQVEAFILRL-----ESAKTNVYVAKGERPAVAVPOSKQVSFATKKELK 295
QY 151 TQNRKMIYVYCNHGHT-----DLKLVAAGVSGMKDIY-----QFLCNAS 191
Db 296 DEIRRSYV--GSSVVFSGDELNKKVSEPPRLRIIGFDFSTLKRMLKPAVFLRPK 353
QY 192 EREVAAFNSCYTADHERAVAAIQ--HWIRGPEAS--DAQLISALRORRRNDVEKIR 245
Db 354 DDEIISGAVFSIHKKLKSLAKIGIAVSRNANPCFAMLAATPSTIHRD--FELPL 412
QY 246 GLMEDTTOLET--DKLALPMSPLSPSPSPISPAKLEN-----282
Db 413 GTF--LVQPTADIRSLP---PINNPISMSNLIETQRLRGMLRSYOGKYNP 466
QY 283 -----SALTVEPSPODKKNGFVDESEPLKCDSTSSG-----SSA 319
Db 467 SLQWVHYKVALALADEIIPD---FVDNLTLYKKAIOKRVGEYMGDVNVAEYRNDI 521
QY 320 LSRNGSEITKEKKDQVLRQVLRDPCDLOPIF--DDMLHFLNP-----ELRVIEEIPQAE 372
Db 522 SDKNGIKKEEEDGPIYKRIKESG-KRIFAEDDRKQLYIGVLDKEIKAL-KVSQLK 579
QY 373 DKL-DRLFETITGVKSOBASQTLDSVSHLPDL 404
Db 580 DILRDGLRVSGKKAD-----LIDNLTNYVKL 607

RESULT 11

S47508
genome polyprotein - plum pox virus (isolate PVV-SC)
N:Contains: coat protein; protein 6K1; protein 6K2; protein CI; protein HCpro; protein N
C:Species: plum pox virus, PPV
A:Variety: isolate PVV-SC
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
A:Accession: S47508
R:Maliss, E.; Debortre, G.; Jelmann, W.; Casper, R.

submitted to the EMBL Data Library, August 1994
A:Description: Complete nucleotide sequence of a plum pox polyvirus isolate (PVV-SC)
A:Reference number: S47508
A:Accession: S47508

A:Molecule type: genomic RNA
A:Residues: 1-3140 (MAI)

A:Cross-references: EMBL:X81083; NID:9531731; PIDN:CAA56974.1; PID:9531732

A:Experimental source: isolate PVV-SC

C:Superfamily: tobacco etch virus genome polyprotein

C:Keywords: ATP; coat protein; genome-linked protein; nucleotide binding; P-loop; rho

F:1-308/Product: protein P1 #status predicted <PI>

F:309-766/Product: protein HCpro #status predicted <HC>

F:767-1116/Product: protein P3 #status predicted <P3>

F:1117-1168/Product: protein 6K1 #status predicted <6K1>

F:1169-1803/Product: protein CI #status predicted <CI>

F:1253-1260/Region: nucleotide-binding motif A (P-loop)

F:1338-1343/Region: nucleotide-binding motif B

F:1342-1345/Region: DEXH motif

F:1804-1856/Product: protein 6K2 #status predicted <6K2>

F:1857-2292/Product: protein N1A #status predicted <N1A>

F:1857-2049/Product: VPg protein #status predicted <VP>

F:2293-2810/Product: protein N1B #status predicted <N1B>

F:2811-3140/Product: coat protein #status predicted <COA>

F:1919/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 5.08; Score 104.5; DB 2: Length 3140;
Best Local Similarity 20.78; Pred. No. 34; Mismatches 102; Indels 95; Gaps 15;
Matches 65; Conservative 52; Mismatches 102; Indels 95; Gaps 15;

QY 92 KHPDINEHLPWMIYVLELLVNVIVVCSIRKSRLLK--GPRDP-----SAIV 139
Db 1024 KMD-----MLVEFLLLSIGATCSMVNEHKLKQLAADRRKKRFRLOYLHRLS 1076
QY 140 EKAGLKSMPTONREKMIYVYCNHGIDILKLVAAQVSGMKDIYQFLCNASEREVAFS 199
Db 1077 EKVG---CTPTAN--EFLEVGGENPDLLKHAEDLIGDG---QVVHOSKRD----- 1120
QY 200 NGYTADEHERAVAAIQHTITNGPEASLAQLISALRORRRND--VYKIRIGLMDTTO-- 253
Db 1121 --SOANLERVVA-----FVALVWMLPDSRSQGVYKILNKLKIGMSVQAV 1165
QY 254 -----LETDKLALP--MSPSPSPS-PIPSPAKLENSALLTVEPSPODKKNGF 299
Db 1166 HHOSLDIEDILDEKTLTVDFVLOSNVAPVPDSTFEKKMTQLETGANVPIRYRGEY 1225
QY 300 FVD-----ESEPLLRCDSTSSGSSA-----LSRNGSFT---TKKKDT 334
Db 1226 FLEPTREMAHIANEVVHGHODILIRGAVGSGSTGLPFLSKGHVLLLETPRPLAEN 1285
QY 335 VLNQVRIDPCDLOP 348
Db 1286 VCKQLRGQPFNVNP 1299

RESULT 12

S51323
SAC3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YD8358.13; protein YDR159w
C:Species: Saccharomyces cerevisiae
C:Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
C:Accession: S51323; S57983; S71744
R:Bauer, A.; Koelling, R.
submitted to the EMBL Data Library, January 1995
A:Description: The SAC3 gene codes for a nuclear protein required for normal mitosis.
A:Reference number: S51323
A:Accession: S51323
A:Molecule type: DNA
A:Residues: 1-1301 (BAU)
A:Cross-references: EMBL:Z47805; NID:9634085; PID:9634086
R:Korpiy, V.; Richards, C.; Harris, D.
submitted to the EMBL Data Library, July 1995
A:Reference number: S57971

Db 842 DMQPSASVPQMED 855

RESULT 15

T28067

hypothetical protein ZK867.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T28067

R:Name: M.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of *C. elegans* cosmid ZK867.

A:Reference number: 220464

A:Accession: T28067

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-494 <NHA>

A:Cross-references: EMBL:U41039; PIDN:AAA82395.1; CESP:ZK867.1

C:Genetics:

A:Gene: CESP:ZK867.1

A:introns: 3/3; 46/1; 203/2; 242/1; 265/3; 339/1; 402/3

Query Match 4.9%; Score 101.5; DB 2; Length 494;

Best Local Similarity 19.2%; Pred. No. 4.5;

Matches 87; Conservative 57; Mismatches 141; Indels 167; Gaps 22;

QY 30 SSARGKED-VNKTLPNQVYVNHQGPPIHRIHLKLLPSMEATGGEKSSPTIKGP----- 81

Db 61 NEAIKADDIIVKKAP--KIYTKDNGP-----FTNGSTPTTSTATPSVITVSS 105

QY 82 ----KRGHPRONLHKHNDINHLPMWIVLFLLVIVVCSIRKSSRTLK--KGP--- 131

Db 106 ALASSNGCHNNNN-----NNH-----AVNNNLRTIMELEDPDYN 139

QY 132 ---RQDPSAIVEKAGLKKSMTT--TONREKWIYVYNGHIDILKLVAAQVGSQMKDIYQFL 187

Db 140 LIAKSAPTPTVYSKIVATHTVTPRSRPTPKDIK-----EILETIAPSVG----- 182

QY 188 CMASEREVAAFSNGCYTADHERATVAALOHWTIRGPEASILAOLIS-----ALROHRRN 238

Db 183 VSETPEEMCLLPKDASSESDRSV-----LISLGFDFGSTLSLNHOOLQ 225

QY 239 DVEYKIRG---LMDDTQLETDKLLALPMSPEPLSPSPISPNAKLENSALTIV----- 288

Db 226 QVARELKGELISIPDTVQSDHSDDFEODSPPPMAIANISTVGGEATLAAITVAAATNASGQ 285

QY 289 ---EPSPODKNKGKGF---FVDESEPLLRCDSYSSGSSALS-----RNGSFITREK 331

Db 286 RGDGTPDSTDTQKGCSPQRELSPESDP-----STSSGDSGSPPKMLHCKECGTLVRRKS 340

QY 332 -----KDTLVROYRLDPCDLQPIFDMHLFLNPEELRVIE----- 367

Db 341 HLPIMHTMSHGYPPLVAAPVEEKPAPEQIPVNASLSH---NELRVISNAICEIKAAQQA 396

QY 368 IPOAEDKIDRLFEELIGVKSQEAQSOTLDSVYS 399

Db 397 TPVVEQALTYIDSRVG--KLEKSELTALNSITYT 427

Search completed: January 4, 2002, 08:38:38
Job time: 121 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:37:27 ; Search time 13.26 Seconds

(without alignments)
1119.853 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSASVRRPKVLSSI.....SQEASQTLDSVSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : SwissProt-39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103.5	4.9	1301	1	SAC3_YEAST
2	101	4.8	342	1	TOMB_PSEAE
3	100	4.8	746	1	DTNA_MOUSE
4	98.5	4.7	3140	1	POLC_PVRA
5	98	4.7	897	1	APGI_YEAST
6	97	4.6	1971	1	MC3A_MOUSE
7	96	4.6	969	1	YEAC_SCHPO
8	95.5	4.6	733	1	RRE_YEAST
9	95.5	4.6	1025	1	MR21_YEAST
10	95.5	4.6	1483	1	CYPI_YEAST
11	95.5	4.6	3140	1	POLC_PVSK
12	95	4.5	707	1	YJ9C_YEAST
13	95	4.5	3141	1	POLC_PVSK
14	94	4.5	249	1	YCA3_CHLVU
15	94	4.5	994	1	CLC1_MOUSE
16	93.5	4.5	1108	1	DBS_HUMAN
17	93	4.4	465	1	SPE4_CAEEL
18	93	4.4	470	1	SELA_MOOTH
19	93	4.4	723	1	GG43_HUMAN
20	92.5	4.4	910	1	YK63_YEAST
21	92.5	4.4	2468	1	MAPB_HUMAN
22	92	4.4	994	1	STYM_NEUCR
23	91.5	4.4	627	1	Y017_RICPR
24	91	4.4	511	1	CP4B_RAT
25	91	4.4	1770	1	RI15_YEAST
26	91	4.4	2109	1	PKS1_ASPPA
27	90.5	4.3	760	1	FXM1_MOUSE
28	90	4.3	712	1	RK3_CHLRE
29	90	4.3	1443	1	E75C_DROME
30	89.5	4.3	593	1	BINI_HUMAN
31	89.5	4.3	1068	1	AF10_MOUSE
32	89.5	4.3	1637	1	MRSF_STRAU
33	89	4.3	409	1	YNP2_YEAST

34	89	4.3	5430	1	ACF7_HUMAN	09upn3 homo sapien
35	88.5	4.2	1332	1	SPT1_YEAST	P35177 saccharomyc
36	88.5	4.2	1390	1	RPOB_MYCA	P47715 mycoplasma
37	88	4.2	501	1	C72M_ARATH	091tm0 arabidopsis
38	88	4.2	688	1	PYS2_PSEAE	006584 pseudomonas
39	88	4.2	836	1	Y167_HUMAN	009490 homo sapien
40	88	4.2	939	1	ST20_YEAST	003497 saccharomyc
41	88	4.2	1176	1	NIR_NEUCR	P38681 neurospora
42	87.5	4.2	622	1	YNA3_CAEEL	P45894 caenorhabd1
43	87.5	4.2	688	1	EFG_MYCPN	P75344 mycoplasma
44	87.5	4.2	736	1	GEPR_RAT	003555 rattus norv
45	87.5	4.2	914	1	PBPA_BACSU	P39793 bacillus su

ALIGNMENTS

```

RESULT 1
SAC3_YEAST
ID SAC3_YEAST STANDARD; PRT; 1301 AA.
AC P46674;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEUCINE PERMEASE TRANSCRIPTIONAL REGULATOR.
GN SAC3 OR LEPI OR YDR159W OR YDR358.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9027306; PubMed=8873450;
RA Bauer A., Koelling R.
RT "Characterization of the SAC3 gene of Saccharomyces cerevisiae."
RL Yeast 12:965-975(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC SRRAIN-5288C / AB972;
RA Murphy L., Richards C., Harris D., Barrell B.G., Rajandream M.A.,
RL Walsh S.V.;
RN Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RC SRRAIN-5288C / DBY939;
RA Stella C.A., Korch C., Ramos E.H., Mattoon J.R.;
RT "Cloning and sequencing of LEPI, a gene associated with leucine
transport."
RL Yeast 11:460-460(1995).
CC -!- FUNCTION: POTENTIAL REGULATOR OF LEUCINE PERMEASE GENE(S)
EXPRESSION. REQUIRED FOR NORMAL MITOSIS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SAC3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47805; CAAB767.1; -
DR EMBL: Z50046; CA90379.1; -
DR EMBL: U35227; AAN79056.1; -
DR SGD: S0002566; SAC3.
KW TRANSCRIPTION regulation; Nuclear protein.
KW SEQUENCE 1301 AA; 149568 MW; 0679DB1673DDACB6 CRC64;

```

Query Match 4.98; Score 103.5; DB 1; Length 1301;
Best Local Similarity 21.8%; Pred. No. 6.1;
Matches 69; Conservative 46; Mismatches 124; Indels 77; Gaps 12;

QY	100	LPRIATVLELLIVLVIVVCSIRKSSSTLTKGPRQDPSALVEKAGLAKSTPTQNEKMTY	155
QY	160	YCNCNGCIDITLKVAAQGSQWKDIYQFLCNSAREVAASNCGYTADHERAYALQHMV	217
Db	438	LPRLMFFGLQHLMLDLIREYALRALSHPTLKKRHPDIPYLE-----NQLFNKQELIE	491
QY	492	FCNYSYIEIINGDAADL-----KTLQYSHKLSSTQ--PLKTYIYLCLEH--RLQKITY	542
Db	218	--TGPASIALQILI-----SALQRHRNDVYEKTRIGLMEDTTQLETDL	259
QY	543	GLINGGEDNLASSVYVADPKKDRIPSIADQSFEMENFQNNYNEKLNQNSVQPIINTSPK	602
Db	260	ALPMSPS--PLSPSPISPAKXLENSALLTVEP--SPQ-----DKKG	298
QY	603	RVATRPHPHPPSQESKOLP--QISQSHTUSITNLLTPQYHGDLSEKQOQIKTYVDGSP	660
Db	299	FVFESEPLRLCDSTSGSSALSRLNGSFYTKKEKDDVLNQVRLDPCDLPIDPDMLEFLN	358
QY	661	FVFDQSANQNSVFAKSKAHMISTTSNGAY--DEKLSSEQDEM-----K	701
Db	359	PEELRVIEETPQAECK	374
QY	702	KEEORIEEEKTQLKKK	717
RESULT	2		
TONB_PSEAE	TONB_PSEAE	STANDARD:	PRT: 342 AA.
ID	AC	051368;	
DT	DT	01-NOV-1997 (Rel. 35, Created)	
DT	DT	20-AUG-2001 (Rel. 40, Last sequence update)	
DT	DT	20-AUG-2001 (Rel. 40, Last annotation update)	
DE	TONB_PROTEIN		
GN	TONB_OR	PA5531.	
OS	Pseudomonas aeruginosa.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		
OC	Pseudomonas.		
OX	NCBI_Taxid=287;		
RA	RA	[1]	
RC	RC	SEQUENCE FROM N.A.	
RC	RC	STRAIN-ATCC 15692 / PA01;	
RX	RX	MEDLINE=20437337; PubMed=10984043;	
RA	RA	Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,	
RA	RA	Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,	
RA	RA	Gardner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,	
RA	RA	Brody L.L., Coulter S.N., Folger K.R., Kes A., Ladby K., Lim R.M.,	
RA	RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,	
RA	RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;	
RT	RT	"Complete genome sequence of Pseudomonas aeruginosa PA01, an	
RT	RT	opportunistic pathogen."	
RL	RL	Nature 406:959-964(2000)."	
CC	CC	-1- FUNCTION: INTERACTS WITH OUTER MEMBRANE RECEPTOR PROTEINS THAT	
CC	CC	CARRY OUT HIGH-AFFINITY BINDING AND ENERGY DEPENDENT UPTAKE INTO	
CC	CC	THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO	
CC	CC	TRANSDUCE ENERGY FROM THE CYTOPLASMIC MEMBRANE TO SPECIFIC ENERGY	
CC	CC	REQUIRING PROCESSES IN THE OUTER MEMBRANE, RESULTING IN THE	
CC	CC	RELEASE INTO THE PERIPLASM OF LIGANDS BOUND BY THESE OUTER	
CC	CC	MEMBRANE PROTEINS (BY SIMILARITY).	
CC	CC	-1- SUBUNIT: THE ACCESSORY PROTEINS EXXB AND EXBD SEEM TO FORM A	
CC	CC	COMPLEX WITH TONB (BY SIMILARITY).	
CC	CC	-1- SUBCELLULAR LOCATION: PERIPLASMIC. ANCHORED TO THE CYTOPLASMIC	
CC	CC	MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE	
CC	CC	PERIPLASM (BY SIMILARITY).	
CC	CC	-1- SIMILARITY: BELONGS TO THE TONB FAMILY.	

```

CC This SW-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U23764; AAB18654.1; -.
DR EMBL: AE004965; AAC08916.1; -.
DR InterPro: IPR002965; P-rich_extensn.
DR InterPro: IPR003538; TonB.
KW Transport; Protein transport; Repeat; Complete proteome.
KW Transmembrane; Signal-anchor; Inner membrane; Periplasmic;
KW DOMAIN 1
FT DOMAIN 87
FT TRANSMEM 88 108
FT DOMAIN 109 342
FT DOMAIN 159 181
FT DOMAIN 190 199
FT DOMAIN 199 199
FT CONFLICT 272
FT CONFLICT 272 G -> R (IN REF. 1).
SQ SEQUENCE 342 AA; 36692 MW; 157F036B39B5152D CRC64;

```

Query Match	4.88;	Score 101;	DB 1;	Length 342;
Best Local Similarity	20.0%;	Pred. No. 1.5;		
Matches	73;	Conservative	40;	Mismatches 122;
				Indels 130;
				Gaps 17;

OY		64	PSMATGEGKSPTEIKGRPKRHPRONLHKHFDITENHLPMWIVLELLVLVIVVCISIRKS	123
Dd		22	PTAAQGDDESILPCVNAQGERP-NLR-----VVDCS----	52
OY		124	SRTLKGPRDPSAIVEKAGLIKSMPTQNREKIWIYYCNGHGIDILKLVAAYGSOMDKI	183
Dd		53	-----GARRREEVAVE-----EVLIPY- AHGSDPEDVPGEPPKSRM----	87
OY		184	YQFCNANSERVAFNSGYTAHDHRAAALOHMTIRP-----EASLAOLISALROH	235
Dd		88	-----WLSSCAAAVMHAIIIGALY-WVMPFPAPELNLGHGLPRTMQVNFOLE	134
OY		236	RNDVVERKIKGLMDTT--OLETDLKALPLMSPSL-SPSPIPSMAKLEMSALLTYEPSP	292
Dd		135	KKAAPTPOPPAAPAEPEPPKIEERK-EDEPKPEVEKPKPKPKPVENAIIPKAKPKP	193
OY		293	QDKKKGFVDSEPELRLCDSTSSGSA-----LSRNGS-----FYIREK	331
Dd		194	EPRK-----PEPEESTEASSQSPSSAAPPATPVGGSTGAQAAPSGSGAPAGLSGSL	249
OY		332	KDTYLKRVRLDDCDLOPIFLDMHFLNP-----ELRVITEETPOAEDKL	375
Dd		250	NDSOIKFKLPMDP---PYRPMAGARGIEGRVKVLTITSDDRIDLDIQVLESVP-----	299
OY		376	DRLFPE	380
Dd		300	SRMED	304

RESULT 3

ID	DTNA_MOUSE	STANDARD:	PRT: 746 AA.
AC	Q99ZNM; Q61498; Q61499; Q9MUW0; Q9MUL9; P97J19; Q9QZ25;		
DT	20-AUG-2001 (Rel. 40, Created)		
DT	20-AUG-2001 (Rel. 40, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	DYSTROBREVIN ALPHA (ALPHA-DYSTROBREVIN).		
DTNA	OR DTN.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sclurognathii; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 2 AND 7).		
RC	TISSUE=Brain;		
XX	MEDLINE=96205975; Pubmed=8631824;		

RA Blake D.J., Nawroztzki R., Peters M.F., Froehner S.C., Davies K.E.;
RT "Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic
RT protein.";
RL J. Biol. Chem. 271:7802-7810(1996).
RN 121
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC SPRAIN-129/SV: TISSUE=Brain;
RX MEDLINE=97224465; PubMed=9119373;
RA Ambrose H.J., Blake D.J., Nawroztzki R.A., Davies K.E.;
RT "Genomic organization of the mouse dystrobrevin gene: comparative
RT analysis with the dystrophin gene.";
RL Genomics 39:359-369(1997).
RN 131
RP SEQUENCE FROM N.A. (ISOFORM 5).
RC TISSUE=Muscle;
RX MEDLINE=98369056; PubMed=9701558;
RA Nawroztzki R., Loh N.Y., Ruegg M.A., Davies K.E., Blake D.J.;
RT "Characterisation of alpha-dystrobrevin in muscle.";
RL J. Cell Sci. 111:2595-2605(1998).
RN 141
RP SEQUENCE FROM N.A. (ISOFORMS 3; 4 AND 6).
RC SPRAIN-C3H: TISSUE=Muscle;
RX MEDLINE=20035753; PubMed=10570976;
RA Enlrg R.E., Malmone M.M.;
RT "Differential expression and developmental regulation of a novel
RL alpha-dystrobrevin isoform in muscle.";
RN 151
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC SPRAIN-C57BL/6J: TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217651;
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotjohori T., Bono H., Kasukawa T., Saito R.,
RA Kadori K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Knehl P., Lewis S., Matsuo Y., Nakaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Faruqi M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki K., Sato K., Schenwabsch C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshbar-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: INVOLVED IN SYNAPSE MATURATION AND REQUIRED FOR NORMAL
CC MUSCLE FUNCTION.
CC -1- SUBUNIT: INTERACTS WITH DYSTROPHIN, UTROPHIN AND SYNPROPHIN.
CC ISOFORMS 5 AND 6 DO NOT INTERACT WITH STINTROPHIN. ISOFORMS 3 AND 4
CC DO NOT INTERACT WITH UTROPHIN.
CC -1- SUBCELLULAR LOCATION: CYTOSOLASMIC.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS: 1 (SHOWN HERE),
CC 2/ALPHA-DYSTROBREVIN-1/ALPHA-DB1, 3/ALPHA-DYSTROBREVIN-2B/ALPHA-
CC DB2B, 4/ALPHA-DYSTROBREVIN-2A/ALPHA-DB2A, 5/ALPHA-DYSTROBREVIN-
CC 3/ALPHA-DB3, 6/ALPHA-DYSTROBREVIN-3/ALPHA-DB3 AND 7; ARE PRODUCED
CC BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, HEART, LUNG AND
CC BRAIN. SARCOLEMA AND NEUROMUSCULAR JUNCTION IN SKELETAL MUSCLE.
CC ISOFORM 2 IS RESTRICTED TO THE NEUROMUSCULAR JUNCTION. ISOFORMS 5
CC AND 6 ARE ONLY EXPRESSED IN MUSCLE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION OF ALPHA-DYSTROBREVIN IS
CC UPREGULATED DURING DIFFERENTIATION, WITH ISOFORMS 2, 5 AND 6
CC EXPRESSED EARLIEST AND ISOFORM 3 AND 4 EXPRESSED LATER.
CC -1- DOMAIN: THE COILED-COIL DOMAIN MEDIATE THE INTERACTION WITH
CC DYSTROPHIN AND UTROPHIN.
CC -1- PTM: PHOSPHORYLATION OF ISOFORM 2 ON TYROSINE KINASE SUBSTRATE
CC DOMAIN PRESENT IN THE C-TERMINUS.

```
CC -!- SIMILARITY: BELONGS TO THE DYSTROPHIN FAMILY. DYSTROBEVIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95326; CAA64518.1; -.
DR DR EMBL; X95327; CAA64519.1; -.
DR DR EMBL; Z79787; CAB02145.1; JOINED.
DR DR EMBL; Z79788; CAB02145.1; JOINED.
DR DR EMBL; Z79789; CAB02145.1; JOINED.
DR DR EMBL; Z79790; CAB02145.1; JOINED.
DR DR EMBL; Z79791; CAB02145.1; JOINED.
DR DR EMBL; Z79792; CAB02145.1; JOINED.
DR DR EMBL; Z79793; CAB02145.1; JOINED.
DR DR EMBL; Z79794; CAB02145.1; JOINED.
DR DR EMBL; Z79795; CAB02145.1; JOINED.
DR DR EMBL; Z79796; CAB02145.1; JOINED.
DR DR EMBL; Z79797; CAB02145.1; JOINED.
DR DR EMBL; AJ009669; CAA08770.1; -.
DR DR EMBL; AF143544; AAD33915.1; -.
DR DR EMBL; AF143543; AAD33914.1; -.
DR DR EMBL; AF143542; AAD33913.1; -.
DR DR EMBL; AK019477; BAB31746.1; -.
DR MGD; MGI:106039; Dtn.
DR InterPro; IPR000345; Cytochrome_c.
DR InterPro; IPR000433; ZnF_ZZ.
DR Pfam; PF00569; ZnF_ZZ.
DR SMART; SMO0291; ZnF_ZZ; 1.
DR KX Synapse; Coiled coil; Zinc-finger; Phosphorylation;
KM Alternative splicing
FT FT 2N_FING 237 284 ZZ-TYPE.
FT FT DOMAIN 397 447 SYNPROPHIN-BINDING REGION.
FT FT 458 557 COILED COIL (POTENTIAL).
FT FT VARSPLIC 334 334 V -> VDMV (IN ISOFORM 5 AND ISOFORM 7).
FT FT VARSPLIC 363 419 MISSING (IN ISOFORM 2 AND ISOFORM 7).
FT FT VARSPLIC 364 371 PKRSEIVE -> DAHGCV (IN ISOFORM 5 AND
ISOFORM 6).
FT FT VARSPLIC 372 746 MISSING (IN ISOFORM 5 AND ISOFORM 6).
FT FT VARSPLIC 559 567 TGGASSPRS -> VSVPYCRS (IN ISOFORM 4 AND
ISOFORM 7).
FT FT VARSPLIC 568 746 MISSING (IN ISOFORM 2).
FT FT VARSPLIC 569 570 TGGASSPRS -> VHHEITPLEERT (IN ISOFORM
3).
FT FT VARSPLIC 571 746 MISSING (IN ISOFORM 3).
FT FT CONFLICT 241 241 V -> L (IN REF. 1; CAA64519).
FT FT CONFLICT 251 251 I -> M (IN REF. 1, 2, 3 AND 4).
FT FT CONFLICT 259 259 Q -> P (IN REF. 4).
FT FT CONFLICT 416 416 D -> N (IN REF. 4).
SQ SEQUENCE 746 AA; 84049 MM; 0F06282718E21679 CRC64;

Query Match 4.8% Score 100; DB 1; Length 746;
Best Local Similarity 25.1%; Pred. No. 5.3;
Matches 68; Conservative 31; Mismatches 102; Indels 70; Gaps 13;
```

CC 277 NAKLENSALLVPEPQDNKGF-----EVDSEPLLRCDD-----STSGSSALSIR 322
DB 581 IRSA--SACPFTHHPDSDLTGVGVGVCFAFQSSRHRLRSLDLVAADSDITMTMSLVAE 638
QY 323 NGSFITEKKQVLRQVRLDPDCLQPIPFDDM 353
DB 639 LNSEVASFTESTV-----DSEFSRQFEDL 663

RESULT 4
POLG_PPVRA STANDARD: PRT: 3140 AA.
AC PI7767;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GENOME POLYPROTEIN (CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);
DE GENOME-LINKED INCLUSION PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
(EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
DE PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
DE COAT PROTEIN (CP)).
OS Plum pox polyvirus (strain Rankovic) (PPV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12214;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89370814; PubMed=2773595;
RA Lain S., Riechmann J.L., Garcia J.A.;
RT "The complete nucleotide sequence of plum pox polyvirus RNA.";
RN Virus Res. 13:157-172(1989).
RP [2]
RX SEQUENCE OF 1778-2342 FROM N.A.
RX MEDLINE=89268456; PubMed=2658302;
RA Garcia J.A., Riechmann J.L., Lain S.;
RT "Proteolytic activity of the plum pox polyvirus N1A-like protein in
RL Escherichia coli.";
RN Virology 170:362-369(1989).
RP [3]
RX SEQUENCE OF 2263-3140 FROM N.A.
RA Lain S., Riechmann J.L., Mendez E., Garcia J.A.;
RT "Nucleotide sequence of the 3' terminal region of plum pox polyvirus
RL RNA.";
RN Virus Res. 10:325-342(1988).
CC -I- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
MAY BE INVOLVED IN REPLICATION.
CC -I- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -I- GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY
PREFERENCES FOR THE AMINO ACIDS IN P6 -P1' THAT VARY WITH THE
SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY)
FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS
THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES
CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
CC -I- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -I- THE VIRAL RNA OF POTYVIRUS IS EXPRESSED AS A SINGLE
POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
INDIVIDUAL PROTEINS.
CC -I- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -I- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -I- SIMILARITY: BELONGS TO THE POTYVIRUS POLYPROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
DR EMBL: M21847; AAA85458.1;
DR EMBL: M26965; AAA47085.1;
DR PIR: A60009; GNVSR.A.
DR MEROPS: C04.001; --
DR MEROPS: C06.001; --
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR001730; Peptidase_C4.
DR InterPro: IPR001456; Peptidase_C6.
DR InterPro: IPR002540; poly_P1.
DR InterPro: IPR001592; poly_coat.
DR InterPro: IPR001205; RNA_pol_P3d.
DR InterPro: IPR00270; DEAD; 1.
DR Pfam: PF00863; Peptidase_C4; 1.
DR Pfam: PF00851; Peptidase_C6; 1.
DR Pfam: PF00767; poly_coat; 1.
DR Pfam: PF01577; poly_P1; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS: PR00966; NIPOTYPYASE.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELICC; 1.
KM Hydrolyase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KM Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 301 N-TERMINAL PROTEIN.
FT CHAIN 302 914 HELPER COMPONENT PROTEINASE.
FT CHAIN 915 ? PROTEIN P3.
FT CHAIN ? 1168 6 KDA PROTEIN 1.
FT CHAIN 1169 1803 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1804 1856 6 KDA PROTEIN 2.
FT CHAIN 1857 ? GENOME-LINKED PROTEIN.
FT CHAIN ? 2292 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2293 2810 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2811 3140 COAT PROTEIN.
FT BINDING 1919 1919 COVALENT LINKAGE OF VIRAL RNA (BY
SIMILARITY).
FT NP_BIND 1253 1260 ATP (POTENTIAL).
FT SEQUENCE 3140 AA; 355577 MW; 5F3DB07982CF3F5 CRC64;

Query Match 4.7%; Score 98.5; DB 1; Length 3140;
Best Local Similarity 20.4%; Pred. No. 48;
Matches 64; Conservative 52; Mismatches 103; Indels 95; Gaps 15;

QY 92 KRPDNEHLPMWIVFLLLVVIVVCSIRKSSRTLRK--GPRDP-----SAIV 139
DB 1024 KNFD-----MLVFSLLSTIGATCNSMINENHKLQLAADREKKRKRLQVLTRLS 1076
QY 140 EKAGLKSMPTPNRKRKIYYCNGHGIDILKLVAAOVGSQMKDIYOFICMASEREVAAS 199
DB 1077 EKVY-----CTPR--ADEFLVYGGEMPDLLKHAEDLDIGD-----QVVYHQSKR 1120
QY 200 NGYTADHEAYVALQHWITRGPEASIALQILSLRQRRND-----VVEKIRGLMEDTQ-- 253
DB 1121 --SQANLEKRVVA-----FVALVMMLFDSERSDGVYKILNKLKLGMSVDQAV 1165
QY 254 -----LETDKALP--MSPSLSPS-PISPNKLNLSALLVPEPQDNKGF 299
DB 1166 QHOSLDDIDIDILDEKKLYIDFVLOSNEVAPYVFPDSFEKMWTFNOLETGNVLPNRTSEH 1225
QY 300 FVD-----ESEPPLRCDSSTSGSSA-----LSRNGSFI-----TREKDDT 334
DB 1226 FLEFPRENAHIANEVMHGSQDILIRGAVSGKSGTGLPFLHSKKGHVALLIPTPRLAEN 1285
QY 335 VLQVRLDPDCLQPI 348
DB 1286 VCKLRGQPFVNPV 1299

RESULT 5
APGL_YEAST

```

ID  APG1_YEAST      STANDARD:      PRT:      897 AA.
AC  P53104;
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DE  AUTOPHAGY SERINE/THREONINE-PROTEIN KINASE APG1 (EC 2.7.1.-).
GN  APG1 OR AUT3 OR YGL180W OR G1615.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=5288C / FY1679;
RX  MEDLINE=97197971; PubMed=9046087;
RA  Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA  Bruschi C.V.;
RT  "Sequencing of a 40.5 kb fragment located on the left arm of
RT  chromosome VII from Saccharomyces cerevisiae.";
RL  Yeast 13:55-64(1997).
RN  [2]
RP  SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS.
RC  STRAIN=42180;
RX  MEDLINE=97368130; PubMed=9224897;
RA  Matsura A., Tsukada M., Wada Y., Ohsumi Y.;
RT  "Applp, a novel protein kinase required for the autophagic process in
RT  Saccharomyces cerevisiae.";
RL  Gene 192:245-250(1997).
CC  -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN THE REGULATION OF THE
CC  AUTOPHAGIC PROCESS, A PROTEIN DEGRADATION PROCESS. FOUND TO BE
CC  ESSENTIAL FOR THE FORMATION OF AUTOPHAGOSOMES. REQUIRED FOR THE
CC  MAINTENANCE OF CELL VIABILITY UNDER STARVATION.
CC  -1- PTM: AUTOPHOSPHORYLATED. MAY BE REGULATED BY A SECOND PROTEIN
CC  KINASE. THE PHOSPHORYLATION STATE MAY PLAY A ROLE IN THE INDUCTION
CC  OF PROTEIN DEGRADATION UPON STARVATION.
CC  -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  APG1/UNC-51/ULK1 SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: X91489; CAA62794.1; -
DR  EMBL: Z72702; CAA6892.1; -
DR  EMBL: D29991; BAA21481.1; -
DR  HSSP: P24941; IAO1.
DR  SGD: S0003148; APG1.
DR  InterPro: IPR000719; Euk_kinase.
DR  InterPro: IPR002290; Ser_thr_kin_actsite.
DR  Pfam: PF00069; pkinase; 1.
DR  SMART: SM00220; S_TKC; 1.
DR  PROSITE: PS00107; PROTEIN_KINASE_APP; 1.
DR  PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR  PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW  Transferase; Serine/threonine-protein kinase; ATP-binding;
KW  Phosphorylation.
KW  DOMAIN
FT  NP_BIND 24 325 PROTEIN KINASE.
FT  BINDING 30 38 ATP (BY SIMILARITY).
FT  ACT_SITE 54 54 ATP (BY SIMILARITY).
FT  MUTAGEN 172 172 BY SIMILARITY.
FT  MUTAGEN 211 211 D->A: FAILS TO COMPLEMENT APG1 MUTANT.
FT  MUTAGEN 237 237 E->R: FAILS TO COMPLEMENT APG1 MUTANT.
SQ  SEQUENCE 897 AA; 101717 MW; 7f4c785aa3a7cc46 CRC64;

```

Query Match 4.7%; Score 98; DB 1; Length 897;
 Best Local Similarity 24.0%; Pred. No. 9.6;
 Matches 49; Conservative 32; Mismatches 79; Indels 44; Gaps 8;

```

OY  228 LISALRQHRNDVVEKIRGMEDTQLETDKALPMSPSPLSPSPNAKLNSALLT 287
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  419 LVSDRSFREFREYVVEK-----KSVEVNSLADENAQAQFNPNPKHPISTQNOANLLN 470
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY  288 VEPSPDKNKGFVDESE--PLRCDSSTSGSSALSNNGSFITREKKDLYL----- 336
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  471 EOPSP--NQOYFQNOGENRLLKATSSSSGSGSRRPSLVRRRLSISSINPSNALSRA 528
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY  337 -----RQVRLDPCOLQPIPDML-HPLNPELRVIEEIPQAE-KLD 376
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  529 LGIASTRLEFGANQQQQQITSSPPYSQTULNSQLFHELTENITLIDHLOHPETLKL 588
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY  377 RLEFIIGVKSQEAQOTLLDSVSH 400
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  589 NT-NIVSILESLAKAFV--VYSV 609
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Query Match 4.6%; Score 97; DB 1; Length 1971;
 Best Local Similarity 21.8%; Pred. No. 33;
 Matches 67; Conservative 45; Mismatches 98; Indels 98; Gaps 17;
 OY 146 KSMPTQNRKRWIYCGHGIDILKLVAAQVSGMKDIYQFLCNASEREVAFGNGYTAD 205
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 DB 405 KSVRQSKRRREMIYSLG--GVSSLELTAIQ-----KNIPDYL--NDRALIEKHFSTIAK 455
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 206 HERAYV-----ALQHTIRPEASLAQLISALQHRNDVVEKIRGMEDT-----Q 253
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 DB 456 QVRVETRRSKKLAVIHF-FDHASAAAR-----KKGKGLHKDVITFMHKK 500
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

OY 234 LETDCLALPM-----SPSPSPSPSPMAKLENSALLTYEESPQDNKNGFE 300
DB 501 ISPSKLEPLKEKLESEASOGIEDSPFQHSPLSKPIYAPAGSLIS -KSPVKRPSLLK 559
OY 301 VD--SEPELRLCDSSSSGSSALSRRGFSITK-----EKKDVIYLRQV 339
DB 560 MHQFEADP---DSGSESGEGL---GSCVSSLSLTLGTVADTSEEKYLRLDQDRIMRQA 613
OY 340 RLDPCLDPIFDMDLHFLN-----PE-----ELRLVETIPOAEDKLDRLFEITG 383
DB 614 RVRKRTDL----DKARAFVOTCPDMCEKERLYLRETRISQLSVEVVP-GTDVYDHAAYNE 668
OY 384 VKSQEASQ 391
DB 669 YSRSSADQ 676

RESULT 7
YEAC_SCHPO STANDARD: PRT: 969 AA.
ID YEAC_SCHPO 014077;
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 20-AUG-2001 (Rel. 40, last annotation update)
DE PUTATIVE ZINC-PROTEASE C2E11.12C (EC 3.4.99.-).
GN SPAC2E11.12C OR SPACUNK4.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX Schizosaccharomycetes.
NX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AL031180; CA20142.1; -
CC DR InterPro: IPR001431; Peptidase_M16.
CC DR Pfam: PF00675; Peptidase_M16; 1.
CC DR ProSite: PS00143; INSULINASE; 1.
CC KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc.
FT METAL 68 68 ZINC (BY SIMILARITY).
FT ACET SITE 71 71 ZINC (BY SIMILARITY).
FT METAL 72 72 ZINC (BY SIMILARITY).
FT METAL 149 149 ZINC (BY SIMILARITY).
FT METAL 149 149 ZINC (BY SIMILARITY).
SQ SEQUENCE 969 AA; 112144 MW; 283AV0BFCF3913 CRC64;

Query Match 4.6%; Score 96; DB 1; Length 969;
Best Local Similarity 19.0%; Pred. No. 15; Mismatches 145; Indels 124; Gaps 14;
Matches 77; Conservative 60;

OY 14 PKVLSIOEGTVPDNTSSRGKEDVKNKTLRN-----LQVNHQGGPHNHILKLL 63
DB 422 PKGIGEVESLRPNPFALAHNSIEKGLNKEKFLYGLDGEDSDSGTIDSLNH----- 477
OY 64 PSMETGGEKSSPTPKPKRGRHRLNKHNDINELPMIVLFLLLVLVIVGSIKRS 123
DB 478 -----KTSSELVLP-----ANEFIPMSLEVERQPVTTKLKLVNLVAN 515
OY 124 SNTLKKGRQDSALVER----- 141

```

```

DB 516 DKFVRLMHRKDDTFWPKANVEINFISPIARRSPKVSSTLTLYRLIEDALGEYSYPASL 575
OY 142 AGIKSMPTQOREKMIYXCHGIDILKTLVAQVSGWKDI-----YQFCNASEREV 195
DB 576 AGLSLSLSS-----TGIIICISGTFDKLHVLEKYYAMMRDLYKVPQREILKNLEDEL 632
OY 196 AAFSNGCYTADHERAYAAIOH---WTIRGEASLAOLISALRORRNDVYEKIRGLM--- 248
DB 633 KDY-----DALEAVYHRSHVLTWLSEPHSGMNAELREAIKVOYQGMDSFIDLLKQNF 686
OY 249 -----EDTQVL--ETDKLALPMSPSPSPSPSPSPAKL-----ENSALLTVERS 291
DB 687 LESLVHGYNTIEDARKNLIESAOKL---IDPKPVFASQLSRKRAIIVPEGNYIKTVVFN 743
OY 292 PODKNGKGFVDESEPLRCDSSTSSGSSALSRRNGSFYTKKKTDLR 337
DB 744 KEKNSAIWYNLO--ISOLDERSG--ALTRLARQIMKEPTSIIR 785

RESULT 8
RRP6_YEAST STANDARD: PRT: 733 AA.
ID RRP6_YEAST 012149;
DT 20-AUG-2001 (Rel. 40, last sequence update)
DT 20-AUG-2001 (Rel. 40, last sequence update)
DT 20-AUG-2001 (Rel. 40, last sequence update)
DE EXOSOME COMPLEX EXONUCLEASE RRP6 (EC 3.1.13.-) (RIBOSOMAL RNA
DE PROCESSING PROTEIN 6).
GN RRP6 OR UNC733 OR YOR001W.
OS Saccharomyces cerevisiae (baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Petersson B., Sterky F., Uhlen M.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sterky F., Uhlen M.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RA MEDLINE=98250784; PubMed=9582370;
RA Briggs M.W., Burkard K.T.D., Butler J.S.;
RA "Rrp6p, the yeast homologue of the human PM-Scl 100-KDa autoantigen,
RA is essential for efficient 5.8 S rRNA 3' end formation.";
RL J. Biol. Chem. 273:13255-13263(1998).
RN [4]
RP CHARACTERIZATION.
RA MEDLINE=99396719; PubMed=10465791;
RA Allmang C., Petfalski E., Podtelejnikov A., Mann M., Tollervey D.,
RA Mitchell P.;
RA "The yeast exosome and human PM-Scl are related complexes of 3'->5'
RA exonucleases.";
RL Genes Dev. 13:2148-2158(1999).
CC -1- FUNCTION: COMPONENT OF THE NUCLEAR EXOSOME 3->5 EXORIBONUCLEASE
CC COMPLEX. REQUIRED FOR THE 3' PROCESSING OF THE 7S PRE-RNA TO THE
CC MATURE 5.8S rRNA. HAS A 3'-5' EXONUCLEASE ACTIVITY.
CC -1- SUBUNIT: COMPONENT OF THE EXOSOME MULTISUBUNIT RIBONUCLEASE COMPLEX
CC COMPOSED OF AT LEAST 11 PROTEINS: RRP4, RRP40, RRP41/SK16, RRP42,
CC RRP43, RRP44/DIS3, RRP45, RRP46, MTR3, CSL4 AND RRP6 (ONLY IN THE
CC NUCLEAR COMPLEX). ALSO ASSOCIATED WITH THE GTPASE RAN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 1 HR23 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```



```
OY 126 TLKGPDPDSALVERAGLTKSMTPTON--REKWIYCNCHGID-----ILKLVAAQ 175
DB 978 YEPMGSDPPTNLAKEAYAOEALNFAMDYRNCMIPEFNINRNNSLEFDYNNVILSTYCLD 1037
OY 176 VGSOMKDIYOFL-----CN--ASEREVAASNGYTADEHAYAAALQHWITRGPEASL 225
DB 1038 IGRH-----SLOFIVCLILIRAKCGPLTGMRESSITNG-----TSSGFNSSV 1079
OY 226 AQLISALRHRNDVYKINGLMEEDTQLETKLALPMSPSPSPSPSPSPSPSPSPSPSPSPSP 285
DB 1080 EDEVDVKKQSSSEL--KRDQFMKDVNLDSGDSLAETILMSRMLLFQKLTQLSKRYVAI 1137
OY 286 LVEPSPDPKNGEFVDESE--PLLRCDSSTSGSS-----ALSRNGSEFTTEKKDTV 335
DB 1138 -----RMKSTGFEVSLDTPSKSKSGSGSSFMGLNKHPRVSNMNGSLADKD--- 1188
OY 336 LROVRLDPCDLPEDMLHFLND 359
DB 1189 -----QLQKC---PVYQDALGFVSP 1205

RESULT 11
POLG_PPVSK STANDARD; PRT; 3140 AA.
AC 084934;
AD 15-DEC-1998 (Rel. 37, Created)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2); GENOME-LINKED PROTEIN (VIG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA) (EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48); COAT PROTEIN (CP)1].
OS Plum pox potyvirus (strain SK 68) (PPV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Potyviridae; Potyvirus.
OX NCBI_TaxID=103927;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167908; PubMed=8122394;
RA Palkovics L., Burgan J., Balazs E.;
RT "Comparative sequence analysis of four complete primary structures of plum pox virus strains."
RL Virus Genes 7:339-347(1993).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY. GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY PREFERENCES FOR THE AMINO ACIDS IN P6 -PI THAT VARY WITH THE SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY) FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
CC -1- PPM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PPM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/)
```

```
CC or send an email to license@isb.ch.
CC -----
DB EMBL: M92280; AAB05823.1; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR001730; Peptidase_C4.
DR InterPro: IPR001456; Peptidase_C6.
DR InterPro: IPR002540; Poly_P1.
DR InterPro: IPR001592; Poly_coat.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR001234; Trypsin.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00863; Peptidase_C4; 1.
DR Pfam: PF00851; Peptidase_C6; 1.
DR Pfam: PF00767; Poly_coat; 1.
DR Pfam: PF01577; Poly_P1; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS: PR00966; NIAPOVPRASE.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELICC; 1.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase; Covalent protein; Polypeptide; Thiol protease; RNA-directed RNA polymerase; ATP-binding.
KW ATP-binding.
FT CHAIN 1 301 N-TERMINAL PROTEIN.
FT CHAIN 302 914 HELPER COMPONENT PROTEINASE.
FT CHAIN 915 ? PROTEIN P3.
FT CHAIN ? 1168 6 KDA PROTEIN 1.
FT CHAIN 1169 1803 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1804 1856 6 KDA PROTEIN 2.
FT CHAIN 1857 ? GENOME-LINKED PROTEIN.
FT CHAIN ? 2292 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2293 2810 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2811 3140 COAT PROTEIN.
FT BINDING 1919 1919 COVALENT LINKAGE OF VIRAL RNA (BY SIMILARITY).
FT NP_BIND 1253 1260 ATP (POTENTIAL).
FT SEQUENCE 3140 AA; 355992 MW; 837A5A692B56436A CRC64;

Query Match 4.6%; Score 95.5; DB 1; Length 3140;
Best Local Similarity 20.18; Pred. No. 81;
Matches 63; Conservative 54; Mismatches 102; Indels 95; Gaps 15;

OY 92 KHPDINEHLPMNIVFLVLVIVVCSIRKSSRTLKK--GPRDP-----SAIV 139
DB 1024 KMPD-----MLIFGLLSIGATCNMINEHKLKQVADREKRRFKRLQVLYTLRL 1076
OY 140 EKAGIKSMPTQNEKRWIYCNCHGIDILKLVAAQVGSOMKDIYOFICNASEREVAAS 199
DB 1077 EKIG--CTPT--ADEFLLEVQGENPDLSKYAEDLIDG---QVVVHOSKRD----- 1120
OY 200 NGYADHERAAALQHWITRGPEASLAQLISALRHRND-----VVEKIRGLMEDTQ-- 253
DB 1121 --SQANIERV-----AAFVALVWMLFEDSESDGVYKTLNKLKGWMSIDQTV 1165
OY 254 -----LETDKLAP--MSPSPSPS-PIPSPAKLENSALTVEPSPDPKNGF 299
DB 1166 HHONLDLEDMLDEKKLVDFVLOSNEVAPVYPRPSTTEKKMTNOLDFGNVIPHRTSGH 1225
OY 300 FVD-----ESEPLLRCDSSTSGSSA-----LSRNGSEFT--TEKKDT 334
DB 1226 FLEFTRENAAHIANEVMHSGHODILIRGAVSGKSTGPHILSKKGHVLLEPTRPALEN 1285
OY 335 VLROVRLDPCDLP 348
DB 1286 VCKQLRGQPFVNP 1299

RESULT 12
VJ9C_YEAST STANDARD; PRT; 707 AA.
ID VJ9C_YEAST
AC P47166;
DT 01-FEB-1996 (Rel. 33, Created)
```

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHETICAL. 81.2 KDA PROTEIN IN NMD5-HOM6 INTERGENIC REGION.
 GN YJRI34C OR J2120.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN NCBI_TaxID=4932;
 RP SEQUENCE FROM N.A.
 RA Rose M., Koetter P., Entian K.D.;
 RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: SOME, TO PARAMITOSINS, AND MYOSINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 249634; CAA89665.1; -
 DR EMBL: 249635; CAA89667.1; -
 DR SGD: S0003895; YJRI34C.
 KM Hypothetical protein, coiled coil.
 FT DOMAIN 122 473 COILED COIL (POTENTIAL).
 FT DOMAIN 594 706 COILED COIL (POTENTIAL).
 SQ SEQUENCE 707 AA; 81171 MW; BAA7DIA7265BF3D9 CRC64;

Query Match 4.5%; Score 95; DB 1; Length 707;
 Best Local Similarity 21.6%; Pred. No. 12;
 Matches 52; Conservative 44; Mismatches 85; Indels 60; Gaps 11;

OY 223 ASLAOLISALRQHRND--VVEKI-----RGIMED--TTQLETKLALPMS- 265
 DB 454 SOLBOKENLAKHQENSNEKTIIDKIPVELTDSLSMEGNIEDWTLPOENSLTSMSTL 513
 OY 266 SPLSPSPSPSPAKLNSALTLVSPSPD-KNKGFVD-----ESEPILRCD 311
 DB 514 GLESDPSPKIPYNESHETICSESOHFDRKNVDSIDDIPEAALALQIRGESNLSLN 573
 OY 312 SFS-----SSSSALSRNGSITR---EKDTYLRQVRLPDCDLOPIFDDMLH--FLNPEE 361
 DB 574 NSIYIRASVQLSNSNGHISAHLYNKLSTELKLEGLSASKEYLDNLLKRTKANDDI 633
 OY 362 LRVIEE-----IPQAEKDLRLPEIIGVKSQSEASQTLDSVSHLPDL 404
 DB 634 LRLLLENKFNENYNNOKODLLKRVEMQSKLETSLQILGSEKTEOYEE--LENDVSDLEEM 691
 OY 405 L 405
 DB 692 M 692

RESULT 13
 POLG_PVVD STANDARD; PRT: 3141 AA.
 AC P13529; Q84929; P89038; -
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GENOME POLYPROTEIN (CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
 DE COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
 DE 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);
 DE GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
 DE (EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
 DE PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
 DE COAT PROTEIN (CP)).
 OS Plum pox polytivirus (strain D) (PPV).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 CC Potyvirus.

OX NCBI_TaxID=12212;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90098790; PubMed=2602121;
 RA Teyheney P.Y., Tavert G., Delbos R., Ravelonandro M., Dunez J.;
 RT "The complete nucleotide sequence of plum pox virus RNA (strain D).";
 RL Nucleic Acids Res. 17:10115-10116(1989).
 RN [2]
 RP SEQUENCE OF 2810-3141 FROM N.A., AND SEQUENCE OF 2812-2828.
 RA Ravelonandro M., Varveri C., Delbos R., Dunez J.;
 RT "Nucleotide sequence of the capsid protein gene of plum pox
 RT polytivirus.";
 RL J. Gen. Virol. 69:1509-1516(1988).
 RN [3]
 RP REVISIONS TO C-TERMINUS.
 RA le Gall O.;
 RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC CATALYTIC ACTIVITY: NUCLEAR INCLUSION PROTEIN A HYDROLASES
 CC GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY
 CC PREFERENCES FOR THE AMINO ACIDS IN PE-PI' THAT VARY WITH THE
 CC SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY)
 CC FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS
 CC THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES
 CC CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUS POLYPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16415; CAA34437.1; -
 DR EMBL: D00298; BAA00210.1; -
 DR PIR: S06929; GNVSPP.
 DR PIR: JA0078; JA0078.
 DR MEROPS: C04.001; -
 DR MEROPS: C06.001; -
 DR MEROPS: S30.001; -
 DR Interpro: IPR001410; DEAD.
 DR Interpro: IPR001650; Helicase_C.
 DR Interpro: IPR001730; Peptidase_C4.
 DR Interpro: IPR001456; Peptidase_C6.
 DR Interpro: IPR001592; Poly_P1.
 DR Interpro: IPR001205; RNA_pol_P3D.
 DR Interpro: IPR001254; Trypsin.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00863; Peptidase_C4; 1.
 DR Pfam: PF00851; Peptidase_C6; 1.
 DR Pfam: PF00767; Poly_coat; 1.
 DR Pfam: PF01577; Poly_P1; 1.
 DR Pfam: PF00680; NIA-dep_RNA_pol; 1.
 DR PRINTS: PR00966; NIAOTYPASE.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR Hydrolase: Transferase; Thiol protease; RNA-directed RNA polymerase;
 KM Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 KM ATP-binding.

```

FT CHAIN 1 301 N-TERMINAL PROTEIN.
FT CHAIN 302 915 HELPER COMPONENT PROTEINASE.
FT CHAIN 916 2 6 KDA. PROTEIN 1.
FT CHAIN 1169 2 6 KDA. PROTEIN 1.
FT CHAIN 1170 1804 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1805 1857 6 KDA. PROTEIN 2.
FT CHAIN 1858 2 2293 GENOME-LINKED PROTEIN.
FT CHAIN 2294 2811 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2812 3141 NUCLEAR INCLUSION PROTEIN B.
FT BINDING 1920 COVALENT LINKAGE OF VIRAL RNA (BY
FT NP_BIND 1254 1261 ARP (POTENTIAL)).
FT SEQUENCE 3141 AA: 355569 MM: 6C3641C404414DBB CRC64;

Query Match 4.5%; Score 95; DB 1; Length 3141;
Best Local Similarity 20.1%; Pred. No. 88;
Matches 61; Conservative 52; Mismatches 102; Indels 88; Gaps 14;

QY 103 MIVLFLVIVVIVCSIRSSRTLK--GPRDP-----SALVEKAGLKSMTP 150
DB 1029 MLIVFGLLSIGATCSNMIENHKLKQLADREDKRRKRLQVLTLSKVG---CTP 1084
QY 151 TQNRKRIYVYCNHGIDILKLVAAOVGSQMKDIYQFLCNASEREVAFAFNGYTDHERAY 210
DB 1085 T--ADEFLVEVGGENPDLKHAEDLIDG-----QVYVHOSKRD-----SQANLERV 1130
QY 211 ALIQHWTIRPEASLAOLISALROHRRND---VVERIKGLMEDTTO-----253
DB 1131 A-----FVALVMMLDFDSEKSDGVYKILNKLKIMGSDVDAVHOSLDDIEDI 1177
QY 254 LETDKLALP--MSPSPLSPS-PIPSPAKLKENSALLVFEPSPOKNGKPFVD-----302
DB 1178 LDEKLLTVDVVLOSNEKAPVFPDSTFEKMTNOLGTGNVIAPIKRTGCHLEFIRENAAH 1237
QY 303 -----ESEPILRCDSSTSGSSA-----LSRNGSFI-----TREKDTVLYROYRLDPCD 345
DB 1238 IANVMHGSODILIRGAVSGSKSTGLPFLSKKGHVLLIEPTPLAENVCKOLRGQPFN 1297
QY 346 LQP 348
DB 1298 VNP 1300

RESULT 14
YCX9_CHLVU STANDARD; PRT; 249 AA.
ID YCX9_CHLVU STANDARD; PRT; 249 AA.
AC 020173;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL 27.9 KDA PROTEIN IN TRNR-MINE INTERGENIC REGION (ORF249).
OS Chlorella vulgaris.
OC Chloroplast.
OC Eukaryota: Viridiplantae: Chlorophyta: Trebouxiophyceae: Chlorellales;
OC Chlorellaceae: Chlorella.
OX NCBI_Taxid=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / TAMIVA;
RA MEDLINE=97303241; PubMed=9159184;
RA Makasugi T., Nagai T., Kapoor M., Sugita M., Ito S.,
RA Tsudzuki J., Yoshinaga K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugita M.,
RT Complete nucleotide sequence of the chloroplast genome from the green
RT alga Chlorella vulgaris: the existence of genes possibly involved in
RT chloroplast division.
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AB001684; BA57948.1;
CC KM Hypothetical protein; Chloroplast.
CC SEQUENCE 249 AA: 27944 MM: 8A89631F981E4AE8 CRC64;

Query Match 4.5%; Score 94; DB 1; Length 249;
Best Local Similarity 26.4%; Pred. No. 3;4;
Matches 46; Conservative 21; Mismatches 51; Indels 56; Gaps 8;

QY 257 DKIALPMSP-----SPLSPS-----IPS-----PNAKLNSALTYEPSPDK-----295
DB 34 DKSFDPHPVAVRSPFGSRPEFLMVPSTVSNALSSESGSSTSPRROESSIPVLQAS 93
QY 296 -----NKGFVDESEPLRCDSSTSGSSALSRSNCSFTKEKKDTVLROVRLDPCDLOPIF 350
DB 94 GVSANKTPOSTESSALHQRKSSS-----ERKTIITRQRPNSGPISPVTPPELLPTL 148
QY 351 DDMHLFNLPELRYEIRPOAEDKIDRLFEITGVKSOA-----SOTLLDSV 397
DB 149 -----EALKLOEALDFAQLKL-----KTRQVRYTINLRGOKLDTI 184

RESULT 15
ID CLC1_MOUSE STANDARD; PRT; 994 AA.
AC 064347;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHLORIDE CHANNEL PROTEIN, SKELETAL MUSCLE (CHLORIDE CHANNEL PROTEIN
DE 1) (CLC-1).
GN CLCN1 OR CLC1.
OS Mus musculus (Mouse).
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA MEDLINE=97468267; PubMed=9321463;
RA Schmeissle V., Antropova O., Gronemeyer M., Wedemeyer N., Jockusch H.,
RA Bartsch J.W.;
RT "The mouse Clc1/myotonia gene: ETN insertion, a variable AATC repeat,
RT and PCR diagnosis of alleles."
RL Mamm. Genome 8:718-725(1997).
RN [2]
RP SEQUENCE OF 350-467 FROM N.A.
RA MEDLINE=92065955; PubMed=1659665;
RA Steimeyer K., Klocke R., Orland C., Gronemeyer M., Jockusch H.,
RA Gruender S., Jentsch T.J.;
RT "Inactivation of muscle chloride channel by transposon insertion in
RT myotonic mice."
RL Nature 354:304-308(1991).
CC -----
CC CLC1 FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSEPITHELIAL TRANSPORT.
CC -----
CC CLC1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
CC CLC1 TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLES.
CC CLC1 SIMILARITY: TO OTHER CHLORIDE CHANNELS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:52 ; Search time 25.5 Seconds
(without alignments)
2323.149 Million cell updates/sec

Title: US-09-733-956-2

Perfect score: 2091

Sequence: 1 MNSTESNSASVPRKVLSSI.....SQEASQTLDSVSHLPDLL 405

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp-plant:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2091	100.0	655	4 075509	075509 homo sapien
2	1898	90.8	655	11 095P05	095P05 mus musculu
3	1537	73.5	630	13 098SM6	098SM6 gallus gall
4	120.5	5.8	1059	5 09VVC6	09VVC6 drosophila
5	119	5.7	2781	4 09V1G2	09V1G2 homo sapien
6	117	5.6	1151	5 09VAV4	09VAV4 drosophila
7	116	5.5	599	5 09U3W0	09U3W0 drosophila
8	115.5	5.5	2183	11 088783	088783 mus musculu
9	115	5.5	927	10 09LFE24	09LFE24 arbidopsis
10	114.5	5.5	290	4 09Y3Z8	09Y3Z8 homo sapien
11	113.5	5.4	412	4 09H5E0	09H5E0 homo sapien
12	113	5.4	599	5 09V680	09V680 drosophila
13	112	5.4	1822	11 035412	035412 rattus norv
14	111	5.3	1138	4 043166	043166 homo sapien
15	111	5.3	1783	4 095321	095321 homo sapien
16	111	5.3	1804	4 09UNU4	09UNU4 homo sapien
17	108.5	5.2	1180	4 092625	092625 homo sapien
18	107	5.1	309	12 089402	089402 paramacium
19	107	5.1	461	5 021004	021004 caenorhabd1

20	107	5.1	500	10 09LV07	09LV07 arbidopsis
21	106.5	5.1	425	3 09C007	09C007 neurospora
22	105.5	5.0	452	10 09SC06	09SC06 arbidopsis
23	105.5	5.0	524	2 09ZK03	09ZK03 helicobacte
24	105.5	5.0	1405	10 050052	050052 arbidopsis
25	104.5	5.0	607	3 094395	094395 schizosacch
26	104.5	5.0	3140	12 084925	084925 plum pox po
27	103	4.9	519	10 09FYW4	09FYW4 lycopersico
28	103	4.9	849	5 077068	077068 trichomonas
29	103	4.9	931	5 096524	096524 trichomonas
30	103	4.9	1231	11 055092	055092 cavia porce
31	103	4.9	1583	4 015045	015045 homo sapien
32	103	4.9	2168	5 09VOM0	09VOM0 drosophila
33	103	4.9	2176	5 046112	046112 drosophila
34	102.5	4.9	1385	5 09M4Z0	09M4Z0 drosophila
35	101.5	4.9	494	5 023654	023654 caenorhabd1
36	101.5	4.9	514	5 010956	010956 caenorhabd1
37	101.5	4.9	527	2 025564	025564 helicobacte
38	101.5	4.9	739	5 09VZ86	09VZ86 drosophila
39	101.5	4.9	919	4 09N0S7	09N0S7 homo sapien
40	101	4.8	16215	5 09NRS3	09NRS3 drosophila
41	100	4.8	498	2 0914Y4	0914Y4 pseudomonas
42	99.5	4.8	872	3 09U5S2	09U5S2 schizosacch
43	99.5	4.8	887	5 09N627	09N627 leishmania
44	99	4.7	604	3 074491	074491 schizosacch
45	99	4.7	895	2 09CLK7	09CLK7 pasteurrella

ALIGNMENTS

RESULT 1
ID 075509 PRELIMINARY; PRT; 655 AA.
AC 075509;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH RECEPTOR-6)).
GN DR6 OR DJ181J13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pan G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,
RT Vincenz G., Aggarwal B.B., Dixit V.M.;
RT "Identification and functional characterization of DR6, a novel death domain-containing TNF receptor."
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Parker A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068688; AAC34583.1; -;
DR EMBL; AL096801; CAB75692.1; -;
DR HSSP; P07174; INGR.
DR InterPro; IPR00488; Death.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00020; TNFR_c6; 4.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

Query Match 100.0%; Score 2091; DB 4; Length 655;
 Best Local Similarity 100.0%; Pred. No. 4,2e-166;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNSTSSNSASVPRVYLSIQEGTVPDNTSSARGKEDVKTLPNLQVYVNHQGPPIHRL 60
 |||
 DB 251 MNSTSSNSASVPRVYLSIQEGTVPDNTSSARGKEDVKTLPNLQVYVNHQGPPIHRL 310
 |||
 QY 61 KLLPMEATGGERKSTPIGPKRGHPRONLHKHFDINELPMWIVLFLLVLVIVYCSI 120
 |||
 DB 311 KLLPMEATGGERKSTPIGPKRGHPRONLHKHFDINELPMWIVLFLLVLVIVYCSI 370
 |||
 QY 121 KKSSTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGVSGW 180
 |||
 DB 371 KKSSTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGVSGW 430
 |||
 QY 181 KDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 240
 |||
 DB 431 KDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 490
 |||
 QY 241 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 300
 |||
 DB 491 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 550
 |||
 QY 301 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 360
 |||
 DB 551 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 610
 |||
 QY 361 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 405
 |||
 DB 611 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 655
 |||

RESULT 2
 O9EPUS PRELIMINARY: PRT: 655 AA.

ID O9EPUS
 AC O9EPUS:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DR6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=KIDNEY;
 RA Isogai D., Ichino M., Yoshinari M., Yamaura A., Kurokawa F.,
 RA Minami M.;
 RT "Mouse DR6: mouse homolog of human TNFR-related death receptor-6
 (DR6).";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF322069; AAC38116.1;
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00531; death.1.
 DR Pfam: PF00020; TNFR_C6; 4.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SMO0005; DEATH; 1.
 DR SMART: SMO0208; TNFR; 4.
 DR PROSITE: PS00017; DEATH_DOMAIN; 1.
 DR PROSITE: PS00552; TNFR_NGFR.1; UNKNOWN.1.
 DR PROSITE: PS00550; TNFR_NGFR.2; 1.
 SQ SEQUENCE 655 AA; 71909 MW; 1A97C1432799E4FB CRC64;

Query Match 90.8%; Score 1988; DB 11; Length 655;
 Best Local Similarity 92.4%; Pred. No. 5,4e-150;
 Matches 375; Conservative 7; Mismatches 22; Indels 2; Gaps 2;

QY 1 MNSTSSNSASVPRVYLSIQEGTVPDNTSSARGKEDVKTLPNLQVYVNHQGPPIHRL 60
 |||
 DB 311 KLLPMEATGGERKSTPIGPKRGHPRONLHKHFDINELPMWIVLFLLVLVIVYCSI 370
 |||
 QY 121 KKSSTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGVSGW 180
 |||
 DB 371 KKSSTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGVSGW 430
 |||
 QY 181 KDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 240
 |||
 DB 431 KDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 490
 |||
 QY 241 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 300
 |||
 DB 491 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 550
 |||
 QY 301 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 360
 |||
 DB 551 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 610
 |||
 QY 361 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 405
 |||
 DB 611 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 655
 |||

DB 251 MNSTSSNSASVPRVYLSIQEGTVPDNTSSARGKEDVKTLPNLQVYVNHQGPPIHRL 310
 |||
 QY 61 KLLP--SMEATGGERKSTPIGPKRGHPRONLHKHFDINELPMWIVLFLLVLVIVYCSI 119
 |||
 DB 311 KLLPSSMEAT--GKSTAIKAPKRGHPRONAKHFDINELPMWIVLFLLVLVIVYCSI 369
 |||
 QY 120 IKKSSTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGVSGW 179
 |||
 DB 370 IKKSSTLLKGGPRDPSAIVEKAGLKKSLTPQONREKWIYYCNGHGIDILKVAAGVSGW 429
 |||
 QY 180 KWDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 239
 |||
 DB 430 KWDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRRNDV 489
 |||
 QY 240 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 299
 |||
 DB 490 VEKIRGLMEDTTOLETDKALPMSPLSPSPISPNAKLNSALTVEPSPQDNKNGCF 549
 |||
 QY 300 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 359
 |||
 DB 550 VDESEPLRCDSSTSGSSALSRRNGSFITKEKKDVTYLRQVRLDPCDLOPIFDDMLHFLNP 609
 |||
 QY 360 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 405
 |||
 DB 610 ELRYVEETIPQAEKDLRLFEITIGVKSQEAASQTLDSVYSHLPDLL 655
 |||

RESULT 3
 O98SM6 PRELIMINARY: PRT: 630 AA.

ID O98SM6
 AC O98SM6:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DEATH RECEPTOR 6 (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC Brigham J.T., Johnson A.L.;
 RA "Expression of DR6 in the ovary."
 RT Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF349908; AAK29666.1;
 RN KW Receptor.
 FT NON-TECTOR
 SQ SEQUENCE 630 AA; 68977 MW; 53C53BD1C4B25567 CRC64;

Query Match 73.5%; Score 1537; DB 13; Length 630;
 Best Local Similarity 73.7%; Pred. No. 6,8e-120;
 Matches 314; Conservative 25; Mismatches 61; Indels 26; Gaps 5;

QY 1 MNSTSSNSASVPRVYLSIQEGTVPDNTSSARGKEDVKTLPNLQVYVNHQGPPIHRL 60
 |||
 DB 210 LMSVFDLSSSPAPRNSNGTAEPTDYNDTSANGVYGAGSLSSAGTAAQAGSYHKHTS 269
 |||
 QY 61 KLL---PSMEATGGERKSTPIGPKRGHPRONLHKHFDINELPMWIVLFLLVLVIVY 117
 |||
 DB 270 QAMGKQPAQAEAGGKSSIPYRPRGRP--NVHOFDINELPMWIVLFLLVLVIVY 327
 |||
 QY 118 CSIRKSSRTLLKGGPRDPSAIVEKAGLKKSMPTQONREKWIYYCNGHGIDILKVAAGV 177
 |||
 DB 328 CSIRKSSRTLLKGGPRDPSAIVEKAGLKKSTPTQONREKWIYYCNGHGIDILKVAAGV 387
 |||
 QY 178 SQMKDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRR 237
 |||
 DB 388 SQMKDIYQFLCNASREVAAFNSGTADHERAYALQHWITRGEPASLAQLISALROHRR 447
 |||
 QY 238 NDVEKIRGLMEDTT-----QLETDKALPMSPLSPSPISPNAK 279
 |||

Db 448 NDVYEKIRGLMEDTTPYQOMQPOWQODCSNDCKLEBGDKLALPVSPSLSPVPTSP--K 505
QY 280 LENSALLTVPSPQDKNKGFVDESEPLRCDSTSSSSSLSRNGSITTEKKDVTYRQY 339
Db 506 PPAAVLTVPSPSEK-KCFVDESEPLRCDSTSSSSSLSRNGSITTEKKDVTYRQY 564
QY 340 RLDPCLQPFDDMLHFLNFEELRVEIEIQAEDKLDRLFEITGVKQEQESQOTLLDSVY 399
Db 565 RLDPCLQPFDDMLHFLNFEELRVEIEIQAEDKLDRLFEITGVKQEQESQOTLLDSVY 624
QY 400 HLPDL 405
Db 625 HLPDL 630
RESULT 4
QYVVC6 PRELIMINARY; PRT; 1059 AA.
AC Q9VVC6; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE CG14060 PROTEIN.
GN CG14060.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chamee C., Pfeiffer B.D.,
RA Man K.H., Doyle E.G., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,
RA Abri J.F., Agdayani A., An H.-J., Andrews-Plamkoc C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA dePablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foutier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jallil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jimali B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., MCPerson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleg J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissengach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2193(2000).
DR EMBL: AE003525; AAF9387.1; -;
DR FlyBase: FBgn003683; CG14060.

SO SEQUENCE 1059 AA; 115746 MW; F8E12EF89F38A18B CRC64;
Query Match 5.8%; Score 120.5; DB 5; Length 1059;
Best Local Similarity 22.2%; Pred. No. 0.21;
Matches 88; Conservative 47; Mismatches 136; Indels 125; Gaps 18;
QY 2 NSTPSSASAVRPKVLSSIOEGYTPDNTSS--ARGKEDVAKTLPNLQVYHNOGPHRHRI 59
Db 570 SASPSPTKAOIREEV--SIDEESTKDSIEDMTYRKEDFTLDLKEVYIPRAEDLEDDSSHA 627
QY 60 L-----KLPSMEATGGE-KSSTPIKG-----PKRGH 85
Db 628 LVSPCFEADAKLIDALEAVDLKIKVLPNDVTYADGERELLANDFEQRTLPSSSF 687
QY 86 PRONLKHHPDINEHL-----PMIVFLVLVYVYVCSIRKSSRTLK-- 128
Db 688 VYQANKROYNNOLLQAVNVTLVNAIEGEOMIMEOKOLMKVRLDELTERDIDQPLKRA 747
QY 129 -KGPDPSPALIVEKAGLKSMTPTQNRKMYCNGHGIDILKLVAAOVGSQWK----- 181
Db 748 EKSPKDDT--LEPESLYEPTIPVMDTK-----TEAQTSSLFRIYTKKS 790
QY 182 ----DIYQFLCNASEREVAFNSNGYTADHERAYALQHWIRGPEASIALQISALROHRR 237
Db 791 ELVEDIYQVTEASTVEVK--PNMWLEGYE-----SIAGQEQACPLTYDGESFAP 839
QY 238 NDVYEK-----RGIMEDTTOLETDKLLALPSPSPSPSPSPSPSPSPSPSPSPSPSP 287
Db 840 EEVVTPTPTSPANGTGLGRNTRDELPLPKPKRVLNKSPILRPAP--PKPAK----- 890
QY 288 VEPSPQDKN-----KGFVDESEPLRCDSTSSSSSS 318
Db 891 --ESEDENIYDTIKGCY--ES--MHCKATSSSSSS 919
RESULT 5
QYVVC6 PRELIMINARY; PRT; 2781 AA.
AC Q9VVC6; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE BROMODOMAIN PHD FINGER TRANSCRIPTION FACTOR.
GN BPRF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jones M.H., Hamana N., Shitane M.;
RT "Identification and characterization BPRF a novel bromodomain
RT transcription factor";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032251; BA89208.1; -;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001487; Bromodomain.
DR InterPro: IPR001965; PHD.
DR Pfam: PF00439; bromodomain; 1.
DR Pfam: PF00628; PHD; 2.
DR PRINTS: PR00503; BROMODOMAIN.
DR PROSITE: PS00633; BROMODOMAIN_1; 1.
DR PROSITE: PS50014; BROMODOMAIN_2; 1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR SMART: SM00297; BROMO; 1.
DR SMART: SM00249; PHD; 2.
SQ SEQUENCE 2781 AA; 311210 MW; 7B3BEDC1C4C3CFBE CRC64;
Query Match 5.7%; Score 119; DB 4; Length 2781;
Best Local Similarity 18.0%; Pred. No. 1;
Matches 79; Conservative 78; Mismatches 169; Indels 112; Gaps 18;

```

OY 15 KVLSSIQESTVNDNTSSARG-KEDVNTKLPNLQVNNHGGPPHNLKLLPSMEATGSK 73
DB 2304 QVLSIQSVVNAIOAOQSGVPOQIKLDP-IOI-QQSSAVQTHQIQWVTVQAASVQE 2360
OY 74 SSTPIKGRKGRPRONLKHNDI-NEHLPMLVLELLLVVIVCSIKRSRTLAKGGR 132
DB 2361 QLOQVO-QLRDQOKKKQOQIETKREH-----TLQASNOSEIIOKV 2402
OY 133 QDPAIVEKAGKSKMTPTQNR-KWIVYNGHGIDILKLVAAQVSGMKDIYQFLCNAS 191
DB 2403 MKNHNVIELKKKKMTPTRENOIMYCN-QVMKIYIDKIDKEKO----- 2449
OY 192 EREVAFNSGYADHERAYALQHTIRGPEASLQALISALROHRRNDVVERIGIMEDT 251
DB 2450 -----AAKRRKREESVEQKRSKONATKLSALLPKHKEQLAEILKK-RALLDKD 2497
OY 252 TQLEF-----DKLAL-----PKMSPLSPSPITSPNKLKENS 283
DB 2498 LQIEVOELKRDILKKEKDLMLQAOAVAAPCPVTVLPAAPAPPPPPPPCVQHT 2557
OY 284 ALLTVEPSPDKNKGFVDESEPLKCDSTSSGSSALSFRNGSFITKE-KKDTVLRQVRD 342
DB 2558 GLLSPTPLPVASOKRREBKO-----SSSKSKKKKMISTSKETKKTILY----- 2604
OY 343 PCDLQPIFDMLHFLN-----PEELRVIEE--IPQAEKDLRLREI--I 382
DB 2605 -CICTPYDESFPYIGCDRCQNMVHGRCVGILOSEALIDEVVCQOSTEDAMVLPPL 2663
OY 383 GVKSOEASTGLDVSYS 400
DB 2664 TEKDEYGLKRVLSLOAH 2681

RESULT 6
O9VAY4 PRELIMINARY: PRT: 1151 AA.
ID O9VAY4:
AC O9VAY4:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG5514 PROTEIN.
GN CG5514
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C.M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Davis A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston K., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splitter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
DR EMBL: AE003763; AAF66763.1;
DR Flybase: FBgn0039560; CG5514.
DR InterPro: IPR002965; P_Rich_extensn.
DR PRINTS: PR01217; PRICEXTENSN.
SQ SEQUENCE 1151 AA; 125450 MW; AF8330378C58C1D1 CRC64;

Query Match 5.6%; Score 117; DB 5; Length 1151;
Best Local Similarity 25.8%; Pred. No. 0.47;
Matches 54; Conservative 29; Mismatches 80; Indels 46; Gaps 9;

OY 167 DILKLVAAQVGSQWK-DIYQFLCNASREVAFAFNSGYADHERA---VVALQHTIRGPE 222
DB 3 DFKTLIAEVKSGVDFDEFRCCLADVTKPAVQNRQVETAVNDPLAKQOMT---PE 59
OY 223 ASILAQLSALRQH-RRNDVVERKIRGIMEDTQLETKLAL-----PM 263
DB 60 TNKVQLRERLRKRLHMDSDVLDK--GVDQIVDQVNPVKVATIEPKIESIVYKLGITPRA 117
OY 264 SPSPLSPSPITSPNKLKENSALLTVEP-----SPDKNKGFVDESE-----P 306
DB 118 RPTMSAPLPPEFGALNLSLLNVTGVLPTDLEQISF-DSDRATVKSSESDDELAP 176
OY 307 LTRCDSTSSGSSALSFRNGSFITKEKKDVT 335
DB 177 GVDEEDTSPSYELVSEKTLAIKEELNV 205

RESULT 7
O9U3W0 PRELIMINARY: PRT: 599 AA.
ID O9U3W0:
AC O9U3W0:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SCAVENGER RECEPTOR-LIKE PROTEIN PRECURSOR.
GN SR-CII OR CG8856.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Pearson A.M., Trigatti B., Xu S., Krieger M.;
RT "Identification of a family of genes related to the Drosophila
RT scavenger receptor dsr-CI.";
RT Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF216642; AAF19660.1;
DR HSSP: P10998; 1YVD
DR Flybase: FBgn0020377; sr-CII.
DR InterPro: IPR000998; MAM.
DR InterPro: IPR001212; Somatomedin_B.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00629; MAM; 1.
DR Pfam: PF01033; Somatomedin_B; 1.

```



```

QY 63 -----LPSHEANGGKSSPTPIKGRGHR-QLNKHFDINELHLMWIVLLELLLV 113
DB 105 SSADISTTASQEPQLSNGENPREPTNGSPRMGLMPSDIRE----- 149
QY 114 VIVVCSIRKSSRTL-KKGRDOPSAIVEKAGLKKSMPTQONEKWTYYCNGHIDILKLV 172
DB 150 -LVRSIHETRTDEALSOQPSKSRANVSLKKESSPSRNSNW-----SEGRRVYKL- 202
QY 173 AAQVGSQWKDIYFLCNASEREVAAFNSNGYTDHERAYALQHWITRGPEASIALISAL 232
DB 203 -----KDSPRF--SYDERE----- 214
QY 233 ROHRNDVVEKIRGLMEDTTOLETKLALPMSPSPLSPSPIP-----SPNAKL 280
DB 215 -----TRKTKAKLKEPRRLSLDSRNSNFSRARSSCSPEOELVTGHRRTTSSVAKL 266
QY 281 ENSALLTVPEPSQDKKGFVDESEPLLRCDSTSSGSSALSINSGSYTTEKK----- 332
DB 267 MGLEVPDEPVTQONRENFCDSPRP-----TSRYEVDLQSRGFDLSLKKMMPAKFPK 320
QY 333 -----DTVLQVRLDPCDLQPIFDDM-----LHFLNPE-ELRYIEIIPQAEKLDRL 378
DB 321 ASPMAQVDAKKNQVKIPDATTITLVGEIGKRLSQLEFKKSEKDLRALQILAMEKTOQL 380
QY 379 FELIGVKSQESQTLDS 396
DB 381 IS-----KDDDDNKTLICSS 394

```

```

RESULT 10
QY9328 PRELIMINARY: PRT: 290 AA.
AC 09Y328;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, last annotation update)
DE HYPOHETICAL.33.3 KDA PROTEIN (FRAGMENT).
GN DKFPP564K112.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Duesterhoef A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049996; CAB43230.1;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 290 AA: 33296 MW; 19298B9DDEDE1DAD CRC64;

```

Query Match 5.5%; Score 114.5; DB 4; Length 290;
 Best Local Similarity 21.1%; Pred. No. 0.12;
 Matches 66; Conservative 46; Mismatches 104; Indels 97; Gaps 13;

```

QY 2 NSTESNSASVPRKYLSSIOEGTVDPDNTSSARGKEDVNKTLPLNLOVNIQOGPRHRIHLK 61
DB 6 SSEDASESLAPRSKYKTIQKDIKEASASKENRDI-----ELSTEE----- 49
QY 62 LIPSEATGEGKSPPIKGRGHRONLKHFDINELHLP-----WIVLLELLLV 113
DB 50 -----KDTEDLDLSSILK-TRKRHKHKEHKGEEVILRLVLSKSEWMDLKEYLALQ 103
QY 114 VIVVCSIRKSSRTL-KGRDOPSAIVEKAGLKKSMPTQONEKWTYYC-----NGHDI 166
DB 104 KASMASLKTITQISESEMETSQVONTGKNE--KTANREE--CRTOEKVATGP 157
QY 167 D-----ILKVAQVGSQWKDIYFLCNASE-----REVAAFNSNG 201
DB 158 QVSGVIVVLIISTEPLRPGKQVRODLAISEVLVYLDLEGGTBCNARFKTPEDAQAVINA 217

```

```

QY 202 YT-----ADHERAYALQHWITRGPEASIALISALROHRNDVVEKIRGL 247
DB 218 YTEINKRHCKWLEILISGHEQAY-----W-----QKTLVROAKLNPREFKRRGT 262
QY 248 MEDTTOLETKL 260
DB 263 EKLTQAEKIRLA 275

```

```

RESULT 11
QY9H5E0 PRELIMINARY: PRT: 412 AA.
AC 09H5E0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE CDNA: FLJ23531 FIS, CLONE LNC60605.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HUMAN LUNG;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Oabayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isigai T., Sugano S.;
RA "NEDO human cDNA sequencing project.";
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK021184; BAB15686.1;
SQ SEQUENCE 412 AA: 46539 MW; D72A6D830BB12B94 CRC64;

```

Query Match 5.4%; Score 113.5; DB 4; Length 412;
 Best Local Similarity 17.9%; Pred. No. 0.23;
 Matches 70; Conservative 59; Mismatches 133; Indels 129; Gaps 12;

```

QY 9 SASVPRKVL-----SIOEGTVDPDNTSSARGKEDVNKTLPLNLOVNIQOGPRHRIHLK 63
DB 94 SLKFRKQLFHRMSLKLKHNRPHNSHPKLOSQSQPSVQSQSVYRQSPQTRIPST 153
QY 64 PSEMATGGE-----KSSPTPIKGRGHR--QLNKHFDINELHLMWIVL 107
DB 154 POLSPGQSQSVQTTSPRIPIQPHSLQIPQCGPQSQPVYVMK----- 199
QY 108 LLLVIVVCSIRKSSRTLKGRDOPSAIVEKAGLKKSMPTQONRE-KWTYYCNGHDI 166
DB 200 -----NAVLENLKKQKSMTPAREENORMIVCN--- 227
QY 167 DILKLVAAQVGSQWKDIYFLCNASEREVAAFNSNGYTDHERAYALQHWITRGPEASIA 226
DB 228 QVMKTIILDKIDEEK-----AKKKRRESVEOKRKNQATKLS 267
QY 227 QLISALRQHRNDVVEKIRGLMEDTTOLETK----- 258
DB 268 ALLPKHKEQLRAELIK-RALLDKDLQLEVGELKRDLIKKEKDLMLQATVAAPCP 326
QY 259 ---LALPMSPSPLSPSPISPAKLNSALLTVPEPSQDKKGFVDESEPLLRCDSTSS 315
DB 327 PVTPAPPPAPPPAPPPSPPPAVQHGGLSTPLPLPAAQKRRKREEEK-----SS 378
QY 316 GSSALSIRNGSFTKE-KRDTVROYRLDPCD 345
DB 379 SKSKKKKMIISTSKETKDTKLYCTICKTPYD 409

```

```

RESULT 12
QY9V680 PRELIMINARY: PRT: 599 AA.
AC 09V680;

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG8856 PROTEIN.
 SR-CIT OR CG8856.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 RX MEDLINE-20196006; PubMed-10731132;
 RC STRAIN-BERKELEY;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amaratilake P.G., Scherer S.E., Li P.W., Hoskins K.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Aghayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
 Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Dugan Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W.,
 Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 Jatalai M., Kalush F., Karpen G.H., Ke Z., Kenson J.A., Ketchum K.A.,
 Kimmel B.E., Kodita C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003823; AAF5851.1; -;
 DR HSSP: P10998; IVVD.
 DR FLYBASE: FBgn0020377; Sr-CIT.
 DR InterPro: IPR000998; NAM.
 DR InterPro: IPR001212; Somatomedin_B.
 DR InterPro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00629; NAM; 1.
 DR Pfam: PF01033; Somatomedin_B; 1.
 DR Pfam: PF00084; sush1; 2.
 DR SMART: SM00032; CCP; 2.
 DR SMART: SM00137; NAM; 1.
 DR SMART: SM00201; SO; 1.
 DR PROSITE: PS00560; NAM_2; 1.
 DR PROSITE: PS00524; SOMATOMEDIN_B; 1.
 SO SEQUENCE 599 AA; 67146 MW; 8A70752A7997EF89 CRC64;

Query Match 5.4%; Score 113; DB 5; Length 599;
 Best Local Similarity 23.9%; Pred. NO. 0.42;
 Matches 55; Conservative 38; Mismatches 61; Indels 76; Gaps 12;

QY 3 STESSASVAPKVLSSIOEG---TVPDMSSARGKEDVNTKLPLNLYVNHQGRPHRH 58
 DB 403 TTTSISKNITTEKVFSLSLASNSFTTPRTSTN-----STTPRQISTK----- 447
 QY 59 ILKLPSMEATGKSSSPIKGRGHP-----RONLHK---HFDIN 97
 DB 448 --KLNTNHE-TLPTSTIPTRNNKLSQPSIFSSQHTVTTTTDSRRKDYHEDIAGOLDMN 504
 QY 98 EHLFW--MYLFLLL-VLYVYVCSIRKSSRTLLKGRDPSAIV--EKA--GLKSKTP 150
 DB 505 TNPFPALIVMYLLIGVILVIVANLKOLCKILSKNSSRDEKVFSEKFAEGLRKPRL 564
 QY 151 TONREKWIYCNCHGIDILKLVAAQVGSQMKDIYQFCNMSEREVAFSN 200
 DB 565 SQN-----CNG-----MDQHLCSAFDEDDLYFED 588
 RESULT 13
 ID 035412 PRELIMINARY; PRT; 1822 AA.
 AC 035412;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SPA-1 LIKE PROTEIN P1294.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RX [1]
 RC SEQUENCE FROM N.A.
 RA Takeuchi M., Ide N., Hata Y., Takai Y.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF026504; AAB81526.1; -;
 DR InterPro: IPR000331; Rap_GAP.
 DR InterPro: IPR001478; PDZ.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF02145; Rap_GAP; 1.
 DR PROSITE: PS0106; PDZ; 1.
 DR SMART: SM00228; PDZ; 1.
 SO SEQUENCE 1822 AA; 201923 MW; 479821CE8820B4FD CRC64;

Query Match 5.4%; Score 112; DB 11; Length 1822;
 Best Local Similarity 20.0%; Pred. NO. 2.2; Mismatches 152; Indels 148; Gaps 19;
 Matches 91; Conservative 65; Mismatches 152; Indels 148; Gaps 19;
 QY 2 NSTESNSASVAPKVLSSIOEGTVPDMSSARG--KEDVNTKLPLNLYVNHQGRPHRH 59
 DB 1429 NSTPISINDATSHSTMS-RISASPVVFSARSRRKELHPTTSS----- 1472
 QY 60 LKLPSMEATGKSSSPIKGRGHPRONLHKH----- 94
 DB 1473 -QLAPSFSS-----SSSSSGSPRTFFPROGATSKYLIGMKKPEGTINSVGFMDTRKHQS 1526
 QY 95 DINEHLPMIVLFLVLLVIVYVCSIRKSSRTLLKGRDPSAIVKAGLKKSM---TPT 151
 DB 1527 DQNE-----IAHTRLRASTRDLASPRPSKSTIE-DLKLTLDESPT 1569
 QY 152 QNREKWIYCNCHGIDILKLVAAQVGSQMKDIYQFCN---ASERVAVAFSNGYTDH- 206
 DB 1570 PESQKNKF---HGLSSPQSPFTPTSPSRALHRTLDESYSSQREHFFTSRSLDDA 1626
 QY 247 LMEPTQLFTDKLALPMSPLSPSPISPRAKLNSALLTVESPDCKKNGFVDSSE- 305
 DB 1685 LM-----PLPDTASDLDWNSLVDAKAVYVQARSFAASDEN 1721
 QY 306 -PLLRCDSTSSGSSAL-----SRNCSFTTKEKKDVLQROVRLDPDLOPIF 350

```
Db 1722 HRPL-----SASNSDQLEQALVOMKSYSSKSDSPTLASKVDQLEGMLKMLREDLKKEK 1777
QY 351 DMLHF-----LNPELRVIEIIPQAEKDLRLFE 380
Db 1778 EDKAHLQAEVQHLREDNLRLOEESQNASDKLKKFTE 1813

RESULT 14
043166 PRELIMINARY: PRT: 1138 AA.
ID 043166:
AC 043166:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE KIA0440 (FRAGMENT).
GN KIA0440.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Ishikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,
Tanaka A., Kotani H., Nomura N., Ohara O., Chiba M., Miyajima N.,
Submitted (Oct-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB007900; BAA23712.1;
DR InterPro; IPR000331; Rap_GAP.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF02145; Rap_GAP; 1.
DR PROSITE; PS50106; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
DR NON_TER 1
SQ SEQUENCE 1138 AA; 126014 MW; CEE33A605029F325 CRC64;

Query Match 5.3%; Score 111; DB 4; Length 1138;
Best Local Similarity 20.2%; Pred. No. 1.5;
Matches 92; Conservative 59; Mismatches 157; Indels 148; Gaps 19;

QY 2 NSTESNSASVPRKVLSTIOEGTVDPDNTSSARG--KEDYNKTLPNLQVNVHQGGHHRHI 59
Db 745 NSTFSINDASHTSTWSS-RHSASPVVFTSARSSPKKEELHAPAPS----- 788
QY 60 LKLPSEMTGEGKSTPIKGRGHPRONLKHNF----- 94
Db 789 -OLAPFS--SSSSSGPRSFYPROGATSKYLIGMKKPEGTINSVGFMDTRKRHQS 842
QY 95 DINEHLPWIVLELLLVVIVVCSIRKSSRTLKGRPDPSAIVKAGLKKSM---TPT 151
Db 843 DQNE-----IAHTRLRASTRDLRASPKPTSKSTIE-DLKLIDLESPT 885
QY 152 QUREKWIYVNGHGIDILKLVAAVGSOMKDIYOFICNASEREVAASNGYADHERAVA 211
Db 886 PESQKSKFR---HALSSPSPFPSTTSRRALHRTLSDESI-----YNSQRHFFPT 933
QY 212 ALQHWITRGPASLAQLISALRQHRNDV-----VEKIRGLM 248
Db 934 S-----RASL-----LDQALPNDVLFSSSTVPSLPKSLPLRRPSYTLGKMSLHGEF 978
QY 249 EDTTQLETKKALPMPSPPLSPS--PIPSPAKLENSALLTVPSPQDNKKGFFVDESE- 305
Db 979 SASDSSLTDTIQETRRQPM-PDGLMPLPTADLDWSNLVDAKAVEVORASFASDEN 1037
QY 306 ---PLLRCDSTSSGSSAL-----SRNGSFITREKKDTVLRQVRLDPCDOLPTE 350
Db 1038 HRPL-----SASNSDQLEQALVOMKSYSSKSDSPTLASKVDQLEGMLKMLREDLKKEK 1093
QY 351 DMLHF-----LNPELRVIEIIPQAEKDLRLFE 380
Db 1094 EDKAHLQAEVQHLREDNLRLOEESQNASDKLKKFTE 1129
```

```
RESULT 15
095321 PRELIMINARY: PRT: 1783 AA.
ID 095321:
AC 095321:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PUTATIVE GAP PROTEIN ALPHA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99078010; PubMed=9858596;
RA Gao Q., Strinivasan S., Boyer S.N., Wazer D.E., Band V.;
RT "The E6 oncoproteins of high-risk papillomaviruses bind to a novel
RT putative GAP protein, E6p1, and target it for degradation.";
RL Mol. Cell. Biol. 19:733-744(1999).
DR EMBL; AF090989; AAD12543.1;
DR InterPro; IPR000331; Rap_GAP.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF02145; Rap_GAP; 1.
DR PROSITE; PS50106; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 1783 AA; 197493 MW; 8C9FB14DBD867FCC CRC64;
```

```
Query Match 5.3%; Score 111; DB 4; Length 1783;
Best Local Similarity 20.2%; Pred. No. 2.6;
Matches 92; Conservative 59; Mismatches 157; Indels 148; Gaps 19;

QY 2 NSTESNSASVPRKVLSTIOEGTVDPDNTSSARG--KEDYNKTLPNLQVNVHQGGHHRHI 59
Db 1390 NSTFSINDASHTSTWSS-RHSASPVVFTSARSSPKKEELHAPAPS----- 1433
QY 60 LKLPSEMTGEGKSTPIKGRGHPRONLKHNF----- 94
Db 1434 -OLAPFS--SSSSSGPRSFYPROGATSKYLIGMKKPEGTINSVGFMDTRKRHQS 1487
QY 95 DINEHLPWIVLELLLVVIVVCSIRKSSRTLKGRPDPSAIVKAGLKKSM---TPT 151
Db 1488 DQNE-----IAHTRLRASTRDLRASPKPTSKSTIE-DLKLIDLESPT 1530
QY 152 QUREKWIYVNGHGIDILKLVAAVGSOMKDIYOFICNASEREVAASNGYADHERAVA 211
Db 1531 PESQKSKFR---HALSSPSPFPSTTSRRALHRTLSDESI-----YNSQRHFFPT 1578
QY 212 ALQHWITRGPASLAQLISALRQHRNDV-----VEKIRGLM 248
Db 1579 S-----RASL-----LDQALPNDVLFSSSTVPSLPKSLPLRRPSYTLGKMSLHGEF 1623
QY 249 EDTTQLETKKALPMPSPPLSPS--PIPSPAKLENSALLTVPSPQDNKKGFFVDESE- 305
Db 1624 SASDSSLTDTIQETRRQPM-PDGLMPLPTADLDWSNLVDAKAVEVORASFASDEN 1682
QY 306 ---PLLRCDSTSSGSSAL-----SRNGSFITREKKDTVLRQVRLDPCDOLPTE 350
Db 1683 HRPL-----SASNSDQLEQALVOMKSYSSKSDSPTLASKVDQLEGMLKMLREDLKKEK 1738
QY 351 DMLHF-----LNPELRVIEIIPQAEKDLRLFE 380
Db 1739 EDKAHLQAEVQHLREDNLRLOEESQNASDKLKKFTE 1774
```

Search completed: January 4, 2002, 08:39:09
Job time: 137 sec

seq_name:	gb_est2:BF681798				
seq_documentation	block: *				
LOCUS	BF681798	1039 bp	mrna	EST	21-DEC-2000
DEFINITION	602117463f1 Soares-mammary_gland_MNLKG Mus musculus cDNA clone				
	IMAGE:346885.5', mRNA sequence.				
ACCESSION	BF681798				
VERSION	BF681798.1	GI:11955693			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus				
REFERENCE	1 (bases 1 to 1039)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue procurement: M. Bento Soares, Ph.D. cDNA library Preparation: Bento Soares and M. Fatima Bonaldo cDNA library Arrayed by: The I.M.A.G.E. Consortium (ULNL)				

```
184 TyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAs 200
|||||
451 TATACAGTTCTTCCAAAGCCAGCGAGAGAGGTGGCGCTTCTCCAA 500
|||||
200 ngLYTThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpT 217
|||||
501 TGGATTACACTGCATCATGAAAGGGCCCTACGGCTCTCGACGACATGGA 550
|||||
217 hrLeaArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 233
|||||
551 CCAATCCCTGGCCCTGAGGCCAGCCTTGCCCACTATGAGCGCTTGCGC 600
|||||
234 GlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAs 250
|||||
601 CAGCAGCGAGCAATGATGATGTGTGGAGAGATGTGGGCTGATGAGAGA 650
|||||
250 pThrThrGlnLeuGlnThrAspLysLeuAlaLeu.PrometSerProSer 266
|||||
651 CACCAGCCAGTTGGAAAGACAAAGTGGCTCTCCCAATGAGCCCGAGT 700
|||||
267 ProLeuSerProSerPro.IleProSerProAsnAlaLysLeuGluAsn 283
|||||
701 CCGCTTAGCCCGAGCCCATGCC...AGTCTTAAGCTGAAACTTGAGAAAT 747
|||||
283 eValAlaLeuLeuThrValGluPro..... 290
|||||
748 CCCATCTCTCTACAGTGGAGCCTCAGCCTGGGACAGAAAGTCTCTCG 797
|||||
291 .....SerProGlnAspLys 295
|||||
798 GGCAGAGATCAAAACCTTTGGAGAGCGATCCCATCAGGGTCTCAGACTGG 847
|||||
295 sAsnLysGlyPhePheValAspGluSerGlu..... 305
|||||
848 AAACGGTCTTATTACAAAAACAGGACAGGTGCGAGAGCCGAGCGG 897
|||||
306 .....ProLeuLeuArgCysAspSerThrSerSerGlySerSerAla 319
|||||
898 TGAATGGCCCATCTTGAACGGTGAACCTGACCCAGACGGGTGACAGACGG 947
|||||
320 .....LeuSerArgAsnGlySerPheIleThrLysG1 330
|||||
948 GGAAGAGCCCAAGAAAGATGTCAACAGAAACCACTTGTAGTACACATAAGA 997
|||||
330 uLysLysAspThrValLeuArgGlnValArg 340
|||||
998 CGACGACAGACGATCCACGCGCTATACGG 1028
|||||
```

seq_name: gb_est1:AU133604

seq_documentation_block: 704 bp mRNA EST 24-OCT-2000

LOCUS AU133604 OVARC1 Homo sapiens cDNA clone OVARC1000267 5', mRNA

DEFINITION sequence.

ACCESSION AU133604

VERSION AU133604.1 GI:10994143

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 704)

AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saïto,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.

TITLE HRI human cDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3951

Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

location/Qualifiers

1. 704

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="OVARC1000267"

/clone_lib="OVARC1"

/tissue.type="ovary, tumor tissue"

/note="Vector: pME18SFL3"

BASE COUNT 176 a 204 c 189 g 132 t 3 others

ORIGIN

alignment_scores:

Quality: 1158.00 Length: 234

Ratio: 5.035 Gaps: 0

Percent Similarity: 98.291 Percent Identity: 97.863

alignment_block:

US-09-733-956-2 x AU133604 ..

Align seg 1/1 to: AU133604 from: 1 to: 704

115 lleValValCysSerIleArgLysSerSerArgThrLeuLysGlyPr 131

3 ATGTGAGTGTGACATCCGGAAGAGCTCGAGACTGTGAAGAGGGGCC 52

131 cArgGlnAspProSerAlaIleValGluLysAlaGlyLeuLysSerM 148

53 TCGGAGGATCCAGTGCATTGTGGAAGACAGGCTGGAAGAAATCCA 102

148 eThrProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHis 164

103 TGACTCCACCCAGACAGCGGGAAGAAATGATCTACTGCAATGGCCAT 152

165 GlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLys 181

153 GGTATCGATATCTGTAAGCTTGTAGACAGCCCAAGTGGAGAGCTGGA 202

181 sAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaP 198

203 AGATATCTATCAGTTCTTTCATATGCCAGTGAGAGGAGGTGCTGCTT 252

198 hSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGln 214

253 TCTCCATGGGTACACAGCCGACAGAGCGGCTTACGACGCTCTGAG 302

215 HisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAl 231

303 CACTGAGCAATCCGGGGCCCGAGCCAGCCTCGCCGCAATTAATTAAGCGC 352

231 alaLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuM 248

353 CCTGCGCCAGACCGGAGAAAGATGTTGTGGAGAAAGATTCTGGGCTGGA 402

248 etGluAspThrThrGlnLeuGlnThrAspLysLeuAlaLeuPrometSer 264

403 TGGAAAGACACACCCAGCTGGAATGCAAAACTGACTCTCCGATGAGC 452

265 ProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG1 281

453 CCAAGCCCGCTTAGCCGAGGCCATCCCAAGCCCAACGGAACCTTGA 502

281 uAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysG 298

503 GAATTCGCTCTCTGAGCGGTGAGCCTTCCCAAGGACAAAGAAAGAGG 552

298 lYPhePheValAspLysSerGluProLeuLeuArgCysAspSerThrSer 314

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 714)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTS: uc09109.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.lnl.gov/db/brp/image/image.html
MGI:1029229
Seq primer: -40RP from Gibco
High quality sequence stop: 426.
Location/Qualifiers
1. 714
/organism="Mus musculus"
/strain="129 - C57/B6 - FVBN"
/db_xref="taxon:10090"
/clone="IMAGE:2648777"
/clone_lib="NCI-CGAP Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab="most-DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."
BASE COUNT 186 a 215 c 184 g 128 t 1 others
ORIGIN
alignment_scores:
Quality: 956.50 Length: 237
Ratio: 4.470 Gaps: 6
Percent Similarity: 90.295 Percent Identity: 84.388
alignment_block:
US-09-733-956-2 x AW211328 ..
Align seg 1/1 to: AW211328 from: 1 to: 714
1 MetAnserThrGluSerAsnSerSerAlaSerValArpProLysValle 17
|||||
13 ATGACATCAACAGATTCACACTACTGCTCTGTATACATAAGTACC 62
17 userSerTleGlnGluGlyThrValProAspAsnThrSerSerAlaarg 34
|||||
63 AAGTGCATCGAGAGGACAGACAGTCCGACATACAGTCAACAGCTG 112
34 LylsGluAspPylAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
|||||
113 GGAAGAGAGGACACTAATAGGACCTGCAACCAACCAACAGTAACCCAC 162
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
|||||
163 CAGCAAGGCCCCACACACAGACATTTGTGAAGCTGCTGCCATGTCAT 212
66 tGluAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysA 83
|||||
213 GGAGGCCACG...GGTGAAGAGTCCAGCAGCAGCATCAAGGCCCCCAAGA 259

83 rgglyHisProArgGlnAsnLysHisIlePheAspIleAsnGluHis 99
|||||
260 GGGGTACCCCGACAGACAGACCTCTCACAGCATTTTGAATCATACAGAGAC 309
100 LeuProTrpMetIleValLeuPheLeuLeuValLeuValIleVal 116
|||||
310 TTGGCCTTGATGATGCTCTCTTTCTTCTGCTTGCCGCGATGATAGT 359
116 lValCysSerTleArgLysSerSerArgThrLeuLysGlyProArg 133
|||||
360 GGTGGCAGTATCCGAAAGAGCTCAGACACTCAAAAAGGGCCCCCGC 409
133 lnaAspSerAlaIleValGluLysAlaGlyLeuLysSerMetThr 149
|||||
410 AGGATCCCGAGCCCTAGTGGAAAGCGGCTGAAGACAGCTCTGACT 459
150 ProThrGlnAsnArgGluLysTrpIleTyTCysAsnGlyHisGlyI 166
|||||
460 CCCAOCAGAAACGGGGAATGATGATCTACCGCCACAGCCCATGTAT 509
166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
|||||
510 TGACATCTTGAAGCTTGTAGCAGCCAGGTGGAAAGCCAGTGAAGAGACA 559
183 letyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
|||||
560 TCTATCATCTTCTTTGCAACCGCACGAGAGAGTGGGCTTCTCC 609
200 AsnGlyTrpThrAlaAspHisGluArgAlaTyAlaAlaLeuGlnHis 215
|||||
610 AATGATCTACTTGAATACAGACAGCGGCTACCGNCTCTGGAGAGAC 659
216 TrpThrTleArgGlyPro...GluAlaSerLeuAlaGlnLeuIleSera 231
|||||
660 CTGGCAGCATTCGTGGCCCTGAGGCCAGCCCTTGCCCGCATTTATACCC 709
231 lalaLeu 232
|||||
710 GCCTG 714
seq_name: gb_est2:BG656232
seq_documentation_block:
LOCUS BG656232 558 bp mRNA EST 05-JUL-2001
DEFINITION l038f11.y1 HR85 1set Homo sapiens cDNA 5' similar to TR:075309
075509 TNFR-RELATED DEATH RECEPTOR-6. ; mRNA sequence.
ACCESSION BG656232
VERSION BG656232.1 GI:13793641
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 558)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Melton D., Brown J., Kenty G., Permutt A., Lee C., Kaestner K., Lemishka I., Seacore M., Brestelli J., Gridwohl G., Clifton S., Hillier L., Werra M., Pape D., Wylie T., Martin J., Blistain A., Schmitt A., Weissing B., Ritzer E., Ronko I., Bennett J., Cardenas M., Gibbons M., McCann R., Cole R., Tsagarelis V., Williams T., Jackson Y., and Bowers Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 475.

FEATURES

Location/Qualifiers
1..558

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by Oligo-dT priming.
Size-selected on agarose gel. Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permut Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 127 a 172 c 148 g 111 t
ORIGIN

alignment_scores: Quality: 953.00 Length: 186
Ratio: 5.124 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.462

alignment_block:
US-09-733-956-2 x BG656232 ..

Align seg 1/1 to: BG656232 from: 1 to: 558

```

181 LYSASRIILEYRGINPHELEUCYSANALASERGLUVALAIAL 197
1  :GGAGATATCATCATCTTTCGAAATGCGAGTGGAGGAGGTGCTGC 50
197 APSESERASNGLYTYRTHRILASPHISGLUARGALATYRALALALEUG 214
51 TTTCTCCAAATGGGTACACAGCCGACGAGCGGGCTTACGACGCTGCG 100
214 LNHSTRPHRTLEARGLYPROGLUALASERLEUALAGLNULESER 230
101 AGCACTGACCATCCGGGGCCCGAGGCCAGCTCGCCCACTAATTAGC 150
231 ALALEARGLINHISARGARGASNPVALIAGLUYSRLEARGLYLE 247
151 GCCCTGGCCGACACCGAGAAACGATGTGTGAGAGATTCGTGGCT 200
247 UMETGLUASPTHRTHRGLEULNTHRASPLYLSLEUALALEUPROMETS 264
201 GATGGAAGACACCCAGCTGGAACACTGACAAACTAGCTCCCGCATGA 250
264 ERPROSERPROLEUSERPROSERPROTLEPROSERPROASNALALYSLEU 280
251 GCCCCAGCCCGCTTAGCCGAGCCCATCCCGAGCCCAACGGGAAACTT 300
281 GLUASSERALALEUETHRVALGLUPROSERPROGLINASPLYSANLY 297
301 GAGAAATTCCTCTCTCTGACGAGTGAGACCTTCCCGACAGACAGACNA 350
297 SGILPHEPHEVALASPLUSERGLUPROLEULEARGCYASPSERTHRS 314
351 GGGCTTCTTGCTGGATGATCGAGGCCCTTCTCCGCTGTGACACTCAT 400
314 ERSERGLYSERSERALALEUSERARGASNGLYSERPHEILETHRLYSGLU 330
401 CCAGCGGCTCTCTCGCGCTGAGCAGAAACGTTCTTTATTACCAAGAA 450
331 LYSGLYASPTHRVALLEUARGINVALIARLEUASPROCYASAPLEUGL 347
451 AAGAAAGACACAGTGTTCGGCAGGTACGCTGACCCCTGTGACTTGCA 500

```

347 nPROTLEPHEASPMETLEUTHISPHELEUASNPGLUGLULEUARGY 364
|||||
501 GCCTATCTTTGATGACATGCTCCACTTTCTTAATCTCGAGAGACTCGGG 550

364 alileglu 366
|||||
551 TGATTGAA 558

seq_name: gb_est2:BG243048

seq_documentation_block:

LOCUS BG243048 861 bp mRNA EST 13-FEB-2001
DEFINITION 602355338P1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4483713 5',
mRNA sequence.
ACCESSION BG243048
VERSION BG243048.1 GI:12752863
KEYWORDS EST.
SOURCE
ORGANISM Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 861)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
CDNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0322 row: k column: 10
High quality sequence stop: 712.

FEATURES

Location/Qualifiers
1..861

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4483713"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 226 a 259 c 221 g 155 t
ORIGIN

alignment_scores: Quality: 942.50 Length: 259
Ratio: 4.116 Gaps: 10
Percent Similarity: 88.417 Percent Identity: 80.695

alignment_block:
US-09-733-956-2 x BG243048 ..

Align seg 1/1 to: BG243048 from: 1 to: 861

```

1 METASERTHRGLUSERASNSERSERALASERVALARGPROLYSVALLE 17
|||||
79 ATGAACTCAACAGATTCACACTACTGCTCTGTGAGAACTAAGTACC 128
17 USERSERILEGLINGLUYTHRVALPROASPASNTHSERSERALARG 34
|||||
129 AAGTGCAATGAGAGAAAGGACAGTGCCTGACAAATACGACTCAACCA 178

```

```

34 1yLysGIuAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
      ||||| :|||:||||| |||||
179 GGAAGGAGGACTAATAGACCTGGCAAAACCAACAAGTTACCCAC 228
      |||||:|||||:|||||:|||||
51 GlnGlnIyProHISHisArgHisIleLeuLysLeuPro...SerMe 66
      |||||:|||||:|||||:|||||
229 CAGCAACCCCCCAGACAGACATCTGAACTGCTGCATGCTCAT 278
      |||||:|||||:|||||:|||||
66 tGAlaIleThcGlyGlyLysSerSerThrProIleLysGlyProLys 83
      |||||:|||||:|||||:|||||
279 GGAAGGACAG...GGTGAAGATCCAGACAGCCATCAAGGCCCCCAAGA 325
      |||||:|||||:|||||:|||||
83 rGcGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlyHis 99
      |||||:|||||:|||||:|||||
326 GGGGTCAACCCAGACAGACAGCCCTCAACAAGCTTTCGACATCAAGAGAC 375
      |||||:|||||:|||||:|||||
100 LeuProTrpMetIleValLeuPheLeuLeuValLeuValIleVal 116
      |||||:|||||:|||||:|||||
376 TTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
      |||||:|||||:|||||:|||||
116 ValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArg 133
      |||||:|||||:|||||:|||||
426 GGTGTGAGATTCGAAAGAGCTCCAGAGCTCAAAAAAGGGCCCGGC 475
      |||||:|||||:|||||:|||||
133 InAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
      |||||:|||||:|||||:|||||
476 AGGATCCAGAGCCATAGTGGAAGAGGGGGGCGGAGAAATCCCTGACT 525
      |||||:|||||:|||||:|||||
150 ProThrGlnAsnArgGluLysTrpIleTyrcysAsnGlyHisGlyI 166
      |||||:|||||:|||||:|||||
526 CCCACCCAGAACCCGAGAAATGATCTACTACCGCAACGCCCATGCTAT 575
      |||||:|||||:|||||:|||||
166 eAspIleLeuLysLeuValAlaIleGlnValGly...SerGlnTrpLysAs 182
      |||||:|||||:|||||:|||||
576 TGACATCTTGAGCTGTAGCAGCCAGGTGGAAAGCCGCTGAGAGGA 625
      |||||:|||||:|||||:|||||
182 rIleTyrcGlnPheLeuCysAsnAlaSerGluArg.GluValAlaAlaPhe 198
      |||||:|||||:|||||:|||||
626 CATCTATCAGTTCTTTGCAACGCCAGAGAGAGGAGTGGCGCTTTC 675
      |||||:|||||:|||||:|||||
199 Ser AsnGlyTyrcThr AlaAspHisGluArgAlaTyrcAla..AlaLeuG 214
      |||||:|||||:|||||:|||||
676 TCCAAATGATACACTTCAGATCATGAAAGGGGCTGAGCGGCTGTC 725
      |||||:|||||:|||||:|||||
214 InHisTrp.ThrIle.ArgGlyProGluAlaSerLeuAlaIleLeuLys 230
      |||||:|||||:|||||:|||||
726 AGCCTGTGAACCATCCGTGGCCCTGAGGCCACCTTGCCCAAGCTCAATT 775
      |||||:|||||:|||||:|||||
230 eAlaIleuArgGlnHisArgArgAsnAspValValGluLysIleArgGly 246
      |||||:|||||:|||||:|||||
776 AAGGCTTGGGCCACCCGAGCAAGAATTGTGTGAAAAAATC...GGG 822
      |||||:|||||:|||||:|||||
247 LeuMetGluAspThr 251
      |||||:|||||:|||||:|||||
823 GCCTTGGAAAAAACCC 837
      |||||:|||||:|||||:|||||
seq_name: gb_est2:BG260064
seq_documentation_block: 589 bp mRNA EST 13-FEB-2001
LOCUS BG260064 602371650F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4479684 5',
DEFINITION mRNA sequence.
ACCESSION BG260064
VERSION BG260064.1 GI:12769880
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 589)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

```

```

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10312 row: c column: 13
High quality sequence start: 7
High quality sequence stop: 589.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC library."
BASE COUNT 146 a 155 c 172 g 116 t
ORIGIN
alignment_scores:
Quality: 925.00 Length: 195
Ratio: 4.894 Gaps: 4
Percent Similarity: 96.923 Percent Identity: 95.385
alignment_block:
US-09-733-956-2 x BG260064
Align seg 1/1 to: BG260064 from: 1 to: 589
53 G1yProHISHisArgHisIleLeuLysLeuLeuProSerMetGluAlaTh 69
|||||:|||||:|||||:|||||
3 GGTCCCAACAGACAGACATCTGAAAGCTGCGCTCATGAGAGGCCAC 52
|||||:|||||:|||||:|||||
69 rGlyGlyLysSerSerThrProIleLysGlyProLys ArgGlyHis 85
|||||:|||||:|||||:|||||
53 TGGGGGCGAGAGTCCAGACGCCCATGAAGGCCCAAGTAGGGGACAT 102
|||||:|||||:|||||:|||||
86 ProArgGlnAsnLeuHisLysHisPheAspIleAsnGlyHisLeuProTr 102
|||||:|||||:|||||:|||||
103 CCTAGACAGAACTACACACAGCATTTTACATCAATGAGCATTTGCCCTG 152
|||||:|||||:|||||:|||||
102 p.MetIleValLeuPheLeuLeuValLeuValIleValIleValCys 118
|||||:|||||:|||||:|||||
153 GTATGATTTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
|||||:|||||:|||||:|||||
119 SerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnAspTr 135
|||||:|||||:|||||:|||||
203 AGTATCCGGAAGAGCTCGAGAGCTCTGAAAAAGGGGCCCGGAGGATCC 252
|||||:|||||:|||||:|||||
135 oSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrProArg 152
|||||:|||||:|||||:|||||
253 CAGTGCCATTTGGAAGAGGAGGCTGAGAAATCCATGATCCCAACCC 302
|||||:|||||:|||||:|||||
152 InAsnArgGluLysTrpIleTyrcysAsnGlyHisGlyIleAspIle 168
|||||:|||||:|||||:|||||
303 AGAACCGGAGAAATGATCTACTCTCAATGGCCATGATGATGATATC 352
|||||:|||||:|||||:|||||
169 LeuLysLeuValAlaIleGlnValGlySerGlnTrpLysAspIleTyrc 185
|||||:|||||:|||||:|||||
353 CTGAGCTTTGAGCAGCCCAAGTGGAGCCAGTGGAAAGATATCTATCA 402
|||||:|||||:|||||:|||||
185 nPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyT 202
|||||:|||||:|||||:|||||

```

```

403 GTTCTTTGGATGCCAGTGAGAGGAGTCTGCTTCTCCAAAGGGT 452
202 yrrhralaasphisgluarqalatyralalaleuqinhistrprrhile 218
|||||
453 ACACAGCCGACGACGAGCGGCGCTACGACACTGTGACGACACTGACCATC 502
219 ArgGlyProGluAlaSerLeu. AlaglinleuIleSerAlaLeuArgGln 224
|||||
503 CGGGGCCCGGAGCGCCGACCCGACCCGACCTAATATGCGCCGCGCCAG 552
235 .HisArgArgAsnAspValValGluLys 243
|||||
553 TGACCCGAGAAACGATGTTGTGAGAAA 580

seq_name: gb_est2:BF682316

seq_documentation_block:
LOCUS BF682316 848 bp mRNA EST 21-DEC-2000
DEFINITION 602117463r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3468855 3', mRNA sequence.
ACCESSION BF682316
VERSION BF682316.1 GI:11956211
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
AUTHORS Email: c9apbs-remail.nih.gov
Tissue Procurement: M. Bento Soares, Ph.D.
Contact: Robert Strausberg, Ph.D.
CDNA Library Preparation: Bento Soares and M. Fatima Bonaldo
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8476 row: m column: 16
High quality sequence start: 21
High quality sequence stop: 715.
Location/Qualifiers
1..848
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_xref="IMAGE:3468855"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 242 a 200 c 232 g 174 t
ORIGIN

alignment_scores:
Quality: 913.00 Length: 207
Ratio: 4.682 Gaps: 1
Percent Similarity: 94.203 Percent Identity: 88.889

alignment_block:
US-09-733-956-2 x BF682316/rev.

Align seg 1/1 to reverse of: BF682316 from: 1 to: 848

```

```

200 AsnGlyTyrThrAlaasphisgluarqalatyralalaleuqinhistr 216
|||||
843 CATGGGTACTCTCCCTCATGACGGCGCTAACGCTGTTCTGACGCTTG 794
216 pThrIleArgGlyProGluAlaSerLeuAlaGlnleuIleSerAlaLeu. 232
|||||
793 GACCATCCGGGCGCTGAGCAACCCCTTGCCAGCTCATTAAGCTCTTGC 744
233 ArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetG1 249
|||||
743 CCAGGCACCGGAGGCAATGATGTTGTGAGAAAGATTCGTGGCTATGGA 694
249 uAspThrThrGlnLeuGluThrAspLysLeuAlaLeuPrometSerPro 266
|||||
693 AGACACACGACGAGTGGAAACAGACAACAGCTGCTCCCATGAGCCCA 644
266 eTrpLeuSerProSerProIleProSerProAlaLysLeuGluAsn 282
|||||
643 GTCCGCTTACCCCGAGCCCATGCCAGTCCCTAAGCTGAACCTTGAGAA 594
283 SerAlaLeuLeuThrValGluProSerProGluAspLysAsnLysGlyPh 299
|||||
593 TCCACTCTCTGACAGTGGAGCCCTCACCGCTGACAGACAGACAGTCTT 544
299 ePheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSer 316
|||||
543 CTTCGTGGAGAGTCAAGAGCCCTTGCGATGCGATCCACATCCAGTG 494
316 LysSerSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLys 332
|||||
493 GCCTCTTCAGCAGTGCAGCAAAAGGCTCTTATTCACAAAGAAAGAAAG 444
333 AspThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProI 349
|||||
443 GACACACTGTGGCGCAGGTCCGCTGAGACCCCTGTGACTTGACGCCAT 394
349 ePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIle 366
|||||
393 CTTTGATGACATGCTGCATATCTGAAACCCCGAGGAGCTGGGCTATTG 344
366 LngLleLeuProGlnAlaGluAspLysLeuAspArgLeuPheGluIle 382
|||||
343 AAGGATTTCCCGAGGCTGAGACAAAGTGAACCGCTTTCGAGATCAT 294
383 GlyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyr 399
|||||
293 GGGGTCAAGAGCCAGAGAGCCAGACCCCTTGTGACTGTGTATACAG 244
399 rHisLeuProAspLeuLeu 405
|||||
243 TCATCTCTCTGACCTATTG 225

seq_name: gb_est1:A1552031

seq_documentation_block:
LOCUS A1552031 548 bp mRNA EST 15-MAR-2000
DEFINITION mm72f10.y1 Stratiogene mouse macrophage (#937306) Mus musculus cDNA
clone IMAGE:533995 5', mRNA sequence.
ACCESSION A1552031
VERSION A1552031.1 GI:4484394
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 548)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Riller
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
TITLE

```

JOURNAL Unpublished (1999)
 COMMENT Contact: Maria M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)

MGI:320931

Seq primer: -40RP from Gibco
 High quality sequence stop: 402

POLYA-No.

FEATURES

SOURCE

Location/Qualifiers
 1. 548
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:533995"
 /clone_lib="Stratagene mouse macrophage (#937306)"
 /tissue_type="macrophage"
 /dev_stage="WEHI-3 cell line"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: blood; Vector: pBluescript SK-; Site:1:
 EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT; WEHI-3 cell line. Average insert size: 1.5 kb.
 3' -3' adaptor sequence: 5' GAATTCGGCAGCAG
 BASE COUNT 131 a 167 c 143 g 107 t
 ORIGIN

alignment_scores:
 Quality: 906.00 Length: 182
 Ratio: 5.061 Gaps: 0
 Percent Similarity: 98.352 Percent Identity: 96.154

alignment_block:

US-09-733-956-2 x A1552031

Align seg 1/1 to: A1552031 from: 1 to: 548

```

148 MethrprothrhgrhlnarGulystRpletyrtyrCysasnGlyH 164
    ::::::::::::::::::::::::::::::::::::::::::::::
2  CTGACTCCACCCAGAACCGGAGAAATGATCTACTACCGCAAGGCCA 51
164 sGlyIleAspIleLeuLysLeuValAlaAlaGlnValGlyserGlnT 181
52 TGGTAATGACATCTTGAAGCTTGAAGCCAGTGGAGAACCCAGTGA 101
181 ySaapIleTyrlGlnPheLeuCySasnaLaserGlnArgGlnValAla 197
102 AGGACATCTATCACTTCTTTCGCAACGCCAGCAGAGAGGAGCTGCC 151
198 PheSerasnGlyTyrlThrAlaAspHisGlnArgAlaTyrAlaAla 214
152 TTCTCCATGATGATACACTGACGATCATGAACGGCCCTACGGCTT 201
214 nHlstrPThrIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSe 231
202 GCACTGGACCATCGTGGCCCTGAGGCCAGCTTGCCAGCTCATTTAC 251
231 lAlaLeuArgGlnHisArgArgAsnAspValValGlnLysIleArgG 247
252 CCTTGGCCGACACCGACGCAATGATGTTGTGAGAAAGATTCTGTC 301
248 MetLysSerThrThrGlnLeuGlnThrAspLysLysIleAlaLeu 264
302 ATGGAAAGACACCGACAGTGTGAAACAGACAACAGGCTCTCCCAT 351
264 rProSerProLeuSerProSerProIleProSerProAsnAlaLys 281

```

```

|||||
352 CCCGAGTCGCTTACGCCGAGCCGACCCATGCCAGTCTACGGAACCTTG 401
281 luanSerAlaLeuLeuThrValGlnProSerProGlnAspLysAsnLys 297
    |||||
402 AGAATTCACATCTCTCTGACAGTGGAGCCCTCACCGCTTGACAGAAC 451
298 GlnPhePheValAspGlnSerGlnProLeuArgGlyCysAspSerThr 314
452 TGTCTTCTTGAGAGAGTCAAGCCCTTCTGCGATCGACTCCACATC 501
314 rSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLys 329
    |||||
502 CAGTGGCTCTTCAGACACTGACAGACAGAAACGGCTCTTTATTCACAA 547
seq_name: gb_estl:AV589747

```

seq_documentation_block:

LOCUS AV589747 575 bp mRNA EST 29-AUG-2000
 DEFINITION AV589747 Bos taurus brain fetus Bos taurus cDNA clone E1BR008C08
 5' mRNA sequence.
 ACCESSION AV589747 GI:9700740
 VERSION AV589747.1
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 575)
 AUTHORS Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jilohzono,A. and
 Suzuki,H.

TITLE bovine cDNA sequencing
 JOURNAL Unpublished (2000)
 COMMENT Contact: Yoshikazu Sugimoto
 Animal Genetics Division
 Shirakawa Institute of Animal Genetics
 Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725
 Email: kazusugi@cocoon.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source

Location/Qualifiers
 1. 575
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone="E1BR008C08"
 /clone_lib="Bos taurus brain fetus"
 /tissue_type="brain"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /note="Vector: pZL1; Site:1: SalI; Site:2: NotI; Poly A
 was deleted from a NotI site"
 BASE COUNT 151 a 166 c 159 g 97 t 2 others
 ORIGIN

alignment_scores:
 Quality: 866.00 Length: 189
 Ratio: 4.838 Gaps: 2
 Percent Similarity: 94.709 Percent Identity: 91.005

alignment_block:

US-09-733-956-2 x AV589747

Align seg 1/1 to: AV589747 from: 1 to: 575

```

1 MetLnsrThrGlnSerAsnSerAlaSerValArgProLysValLe 17
    ::::::::::::::::::::::::::::::::::::::::::::::
13 ATCAACGTCACAGAAATCCAACTCTTCTGCTTGAACCAAAAGTCC 62
17 userSerIleGlnGlyThrValProAspAsnThrSerSerAlaArg 34

```



```

|||||
683 GGACATCATCAGTTCTTGGCAACGCCGACGAGGGAGGTGGCGAC 732
197 apheSeranglyTyThrAlaAspHis 206
733 TTTCTCCAAATGATACATCAGATCAT 760

```

seq_name: gb_est1:BE478092

seq_documentation_block:

LOCUS BE478092 521 bp mRNA 28-AUG-2000
DEFINITION 162175 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE478092
VERSION BE478092.1 GI:9597625

KEYWORDS EST.

SOURCE

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS 1 (bases 1 to 521)
Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and
Wells,K.D.

TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary
gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@psi.barc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGAGC
Plate: 5 row: P column: 6
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..521
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

BASE COUNT 135 a 167 c 135 g 84 t
ORIGIN

alignment_scores:

Quality: 697.00 Length: 151
Ratio: 4.840 Gaps: 0
Percent Similarity: 95.364 Percent Identity: 90.728

alignment_block:

US-09-733-956-2 x BE478092 ..

Align seg 1/1 to: BE478092 from: 1 to: 521

```

1 MetAnserThrgluSerAsnSerSerAlaSerValArgProLysValle 17
:::|||||
68 ATCATACTCAACAGATCCAACTCTTCTGCGCTCTTAGACCAAGGTC 117
17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
||||:|||||
118 AAGTGCACCCAGAGGAGGACAGTCCTGACACACAAAGCTCAGCAGCA 167

```

```

34 lYlYSGluAspValAsnLysThrIleuProAsnLeuGlnValValAsnHis 50
:::|||||
168 GGGAGGAGGGCGGTGACACAGACCTTCCCAAGCTCCAGGTACCAACAC 217
51 GlnGlnGlyProHisLysArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||
218 CAGCAAGGGCCCCACACAGACATCCTGTAAGCTGTGCGGTCATGGA 267
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 84
|||||
268 GGGCAGTGGGGGTGAGAAGTCACACAGCCCATCAAAAGCCCTCAAGAGG 317
84 lYHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
318 GCCACCCAGGAGACACCCACACAGACATTTTGACATCAACGACACCTT 367
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValIleVal 117
|||||
368 CCGTGATGATGTGCTTTTCCCTGCTGCTGCTGCTGCTGATAGTGT 417
117 lYsSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
|||||
418 GTGCAGTATCCGGAAGCTCGAGACTCTCAAGAGGGGCCCGCGCAGG 467
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
468 ATCCACAGCGCATTTGTGAAAAGCAGGAGGCTCAAGAAATCATGAGCC 517
151 Thr 151
|||
518 ACC 520

```

seq_name: gb_est1:AA072902

seq_documentation_block:

LOCUS AA072902 398 bp mRNA 07-FEB-1997
DEFINITION mm72F10.r1 Stralagene mouse macrophage (#937306) Mus musculus cDNA
clone IMAGE:533995 5', mRNA sequence.

ACCESSION AA072902
VERSION AA072902.1 GI:1594632

KEYWORDS

SOURCE

ORGANISM

Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 398)

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNL: contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev1 ET from Amersham

MGI:320931

High quality sequence stop: 341.

location/Qualifiers

FEATURES

source

1..398
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:533995"
/tissue_type="Stralagene mouse macrophage (#937306)"
/dev_stage="WEHI-3 cell line"


```
|||||
435 CAAGGCTCTCTCGTGGATGAGTCGGAGCCACCCTGTCGCTAGTGAC 484
313 Thr.....SerSerGlySerSerAlaLeuSerArgAsn..GlySerPh 326
    ::::::::::::::::::::||
485 TCTCACATCCAGACGGCTCGCTCGCTGAGCAAGGACCGGCTCCTT 534
326 eIleThr.LysGluLysLysAspThrValLeu.ArgGlnValArgLeuAs 342
    ||||| ||||| ||||| ||||| ::|||
535 TATTACCCAAAGAAAGAGACACAGTTGTGGCGGCAAGTG..... 576
342 pProCyAspLeuGlnProIlePheAspAspMetLeuHis..... 355
    :::::: ||::::
577 .....ACGCGCTGGGAACCCATGCTGATTTGAAGTCTTA 610
356 .....PheLeuAsnProGluGluLeuArg 363
    ||||| ||||| ::|||
611 TCTTTGAATGACAATGCTATACATATTCTTAATTCTCTGAGAGATGCGG 660
```

THIS PAGE BLANK (USPTO)

OM of: US-09-733-956-2 to: Issued_Patents_NA.* out_format : pfs

Date: Jan 4, 2002 8:42 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framed+ p2n.model -DEV=xip
-Q/cgcn2_1/USPTO.spool/US0973356/runat_04012002_084025_15531/app_query.fasta.1.467
-DB=Issued_Patents_NA -OPM=fastap -SUFFIX=rni -GAPOP=12.000
-OGAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-OGAPEXT=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bls62
-TRANS=umana40 cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US0973356.ecgcn1_1_80 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPY
-WAIT -THREADS=1
```

Search information block:

```
Query: US-09-733-956-2
Query length: 405
Database: Issued_Patents_NA.*
Database sequences: 351203
Search time (sec): 88.950000
```

Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	1.7e-225	2186
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-959-382-1 +			2091.00	4185.82	9.8e-206	1719	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-042-785A-5 +			1914.00	3830.95	9.8e-205	3331	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-042-785A-1 +			1914.00	3822.29	4.3e-162	1815	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-042-785A-24 +			1527.00	3047.46	8.0e-162	2638	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-042-785A-22 +			1527.00	3042.56	8.0e-162	2638	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-042-785A-3 +			1269.00	2520.83	9.2e-133	2652	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-746-111-4 +			113.50	173.53	0.0397	6585	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-09-041-886-1 +			113.00	179.19	0.0249	3386	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-602-791-1 +			112.00	186.99	0.0091	1600	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-985-950-5 +			107.50	180.82	0.0022	1279	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-985-950-7 +			107.50	176.87	0.0334	1728	
/cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-961-083-55 +			105.00	167.58	0.1102	2389	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-985-916-15 +			101.00	154.36	0.6005	3533	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-324-867-1 +			99.00	141.30	3.21	7032	
/cgcn2_6/ptodata/2/ina/5B.COMB.seq:US-08-989-925-2 +			96.50	153.06	0.7092	1947	
/cgcn2_6/ptodata/2/ina/5B.COMB.seq:US-08-208-887A-48 +			92.50	102.21	481.96	50937	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-315-793-51 +			91.50	136.11	6.24	3282	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-229-515A-12 -			91.00	143.58	2.39	1717	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-645-865-12 -			91.00	143.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-07-872-644-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-297-510-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-479-532-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-455-526-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-455-525-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-09-139-491-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-09-139-491-44 -			91.00	136.08	6.26	3044	
/cgcn2_6/ptodata/2/ina/5A.COMB.seq:US-08-928-941D-28 +			91.00	133.29	8.95	3767	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-928-941D-28 +			91.00	133.29	8.95	3767	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-928-941D-28 +			91.00	133.29	8.95	3767	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-280-590A-30 +			90.00	127.09	19.83	5183	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-280-590A-30 +			90.00	127.09	19.83	5183	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-399-411-3 +			90.00	125.46	24.43	5868	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-399-411-3 +			90.00	125.46	24.43	5868	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-103-840A-1 +			89.50	37.71	6.5e+05	4403765	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-103-840A-1 +			89.50	37.71	6.5e+05	4411529	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-103-840A-1 +			89.50	37.71	6.5e+05	4411529	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-928-941D-2 +			89.00	132.66	9.71	2903	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-928-941D-2 +			89.00	132.66	9.71	2903	
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-08-928-941D-3 +			89.00	132.66	9.71	2903	

```
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-280-590A-2 + 89.00 132.66 9.71 2903
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-280-590A-3 + 89.00 132.66 9.71 2903
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-280-590A-4 - 89.00 130.88 12.20 3325
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-413-574-1 + 88.50 140.11 3.74 1522
/cgcn2_6/ptodata/2/ina/6B.COMB.seq:US-09-413-574-2 - 88.50 135.07 7.12 2235
```

seq_name: /cgcn2_6/ptodata/2/ina/6A.COMB.seq:US-08-959-382-1

```
seq_documentation_block:
; Sequence 1, Application US/08959382
; Patent No. 6013476
; GENERAL INFORMATION:
; APPLICANT: DEEN, KEITH
; APPLICANT: HURLE, MARK
; APPLICANT: YOUNG, PETER
; APPLICANT: TAN, K.B.
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/959,382
; FILING DATE: 28-OCT-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,796
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-959-382-1
```

```
alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-733-956-2 x US-08-959-382-1 ..
Align seq 1/1 to: US-08-959-382-1 from: 1 to: 2186
```

```
1 MetanserthrghinsraserSerAlaSerValArgProLysValle 17
|||||
757 ATGACACCAAGACATCCACTCTTCCCTCTGTAGACCAAGACT 806
|||||
17 UserSer1eIngluclgthrrValProAspAsnThrSerSerAlaArg 34
|||||
807 GAGTACGATCCAGAGGAGGACATCCCTCGACACCAAGCTCAGCAAGG 856
```

```

34 1LYSGIUAAPVALASNYSTRLEUPROASNLEUGLNVAlVALASNHIS 50
    |||||||
857 GGAGGAAGACGTGAAAGACCTCCCAACCTTCAGGTATCAACAC 906
51 GINGINGLYPROHISARGHISILEULYSLLEULEUPROSERMETG 67
    |||||||
907 CAGCAAGGCCCCCAGCACACACATCTGAAAGCTGCTCCATGAGA 956
67 UAlAthrGLyGLyULySserSerThrProLLeYsgLYProlYsArg 84
957 GGCCACTGGGGCGAGAACTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1006
84 LYHISProArgGLINASLeuHISLYSHISPhEASpLLeuSngLHISLeu 100
    |||||||
1007 GACATCTTAGACAGAACCTACACAGCATTTTGACATCAATGACATTTG 1056
101 ProTPMetLLeValLeuPhEuleuLeuValLeuValLLeValVa 117
    |||||||
1057 CCTCGATGATGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
117 LCysSerLLeArgLYSserSerArgThrLeuLYSGLYProArgGLIN 134
    |||||||
1107 GTGAGATATCCGGAAGCTCGAGAGCTGAAAAGGGGGCCCCGAGAG 1156
134 sPProSerAlLLeValGLYSLAGLYLeuLYSserMetThrPro 150
    |||||||
1157 ATCCAGTCCCATTTGTGGAAGGCGAGGCTCAGAAATCCATGATGCTCA 1206
151 ThrGLINASArgGLULySTrPLeTYrTYrCYsASngLYHISGLYLeAs 167
    |||||||
1207 ACCCAGAACCGGAGAAATGATCTACTGCAATGCGCATGATGATCA 1256
167 PLeuLeuLYSLeuValAlAlAGLNVAlGLYserGINTPLYSASPLeT 184
    |||||||
1257 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1306
184 YrGLINPhEuleuCYsASnAlASerGLUArgGLUValAlAlAphESerAsn 200
    |||||||
1307 ATCAGTTCTTTGCAATGCCAGAGAGAGGAGTTCCTGCTTCTTCAT 1356
201 GLYTYrThAlASpHISGLUArgAlATYrAlAlAlALeuGLINHISTrph 217
    |||||||
1357 GGGTACACAGCGACGACGAGGGGCTTACGAGCTCTCAGCACAGGAG 1406
217 rLLeArgGLYProGLUAlASerLeuAlAGLInLeuLLeSerAlALeuArgG 234
    |||||||
1407 CATCCGGGGCCCCGAGGCGACCTCGCCAGCTAATTAGCGCCCTCGCC 1456
234 LNHISArgHrASnASpValValGLULySLLeArgLYLeuMetGLUASP 250
    |||||||
1457 AGCACGCGAGAAACGATGTGTGAGAGATTCGTGGCTGATGAGAAAGC 1506
251 ThrThGLINLeuGLUThrASpLYSLeuAlAlALeuProMetSerProSerPr 267
    |||||||
1507 ACCACCCAGCTGAAACTGAACTAGCTCTCCGATGAGAGCCCCGAGCCC 1556
267 oLeuSerProSerProLLeProSerProASnAlALySLeuGLUASerA 284
    |||||||
1557 GCTTAGCCCGAGGCCCATCCCAAGGCCCAAGCGAATTTGAGAAATTCCG 1606
284 lAlEuleuThrValGLUProSerProGLINASpLYSAsnLYSGLYPhePhe 300
    |||||||
1607 CTCTCTGACGGTGAGACCTTCCCCACAGGACAAAGAAAGGGCTTCTTC 1656
301 ValASpGLUSerGLUProLeuLeuArgCYsASpSerThSerSerGLYse 317
    |||||||
1657 GTGGATGATCGGAGCCCTTCTCCGCTGACTTACATCCAGCGGCTC 1706
317 rSerAlAlASerArgASngLYSerPheLLeThrLYSGLULySAspT 334
    |||||||
1707 CTCGCCGCTGAGAGGAGGTTCTTTATTTACCAAGAAAGAGAGGACA 1756

```

```

334 hrValLeuArgGLINValArgLeuAspProCYsASpLeuGLINProLLePhe 350
    |||||||
1757 CAGTGTGGGCGAGGATGAGCCCTGGAACCCCTGTGACTTGACGCTATCTTT 1806
351 AspAspMetLeuHISPhEuleuASnProGLUGLULeuArgValLLeGLUL 367
    |||||||
1807 GATGACATCTCCACCTTCTTAATCTGAGAGCTGCGGGGTATTGAGAGA 1856
367 uLLeProGLINAlAGLUAAspLYSLeuASpArgLeuPheGLULLeLGLY 384
    |||||||
1857 GATTCGCCAGCTGAGAGACAACTAGACCGGCTATTTCGAATTTATGGAG 1906
384 aLYSserGLINAlASerGLINThrLeuLeuASpSerValTYrSerHIS 400
    |||||||
1907 TCAAGAGCCAGGAGGACGACGACCCCTCTGTGACTCTGTTATATAGCAT 1956
401 LeuProASpLeuLeu 405
    |||||||
1957 CTTCCTGACCTGCTG 1971

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-042-785A-5
seq_documentation_block:
: Sequence 5, Application US/09042785A
: Patent No. 6194151
: GENERAL INFORMATION:
: APPLICANT: Busfield, Samantha J
: TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
: TITLE OF INVENTION: AND USES THEREFOR
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/042.785A
: FILING DATE: 17-MAR-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/938,896
: FILING DATE: 26-SEP-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragoras, Amy E
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: MEI-001CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)742-4214
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1719 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1719
: US-09-042-785A-5

```

```

alignment_scores:
    Quality: 1914.00      Length: 406
    Ratio: 4.883          Gaps: 2
    Percent Similarity: 96.552      Percent Identity: 92.857

```


MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 344..2065
 US-09-042-785A-1

alignment_scores:
 Quality: 1914.00 Length: 406
 Ratio: 4.883 Gaps: 2
 Percent Similarity: 96.552 Percent Identity: 92.857

alignment_block:
 US-09-733-956-2 x US-09-042-785A-1 ..

Align seg 1/1 to: US-09-042-785A-1 from: 1 to: 3331

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProIleValLeu 17
|||||
848 ATGAACTCAACAGATTCCAACTACTGCTCTGTAGAACTAGACCTACC 897
17 userSerIleGluGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
898 AAGTGCATCGAGGAAGGACAGCTGCTGACATACGAGCTCAACCACTG 947
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||
948 GGAAGGAAGGCACTAATAGACCTGCAAAACCAACCAAGTTACCCAC 997
51 GlnGluGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
|||||
998 CAGCAAGCCCCCAACCAACCAACCACTTCAAGCTGCTGCCATCTCCAT 1047
66 tGluAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysA 83
|||||
1048 GAGAGGCCAGC...GGTGAAGATCCACGACAGCCATCAAGCCCCCAAGA 1094
83 rGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
|||||
1095 GGGGTCAACCCACAGAACGCTCAACAGCATTCGACATCAACGAGCAC 1144
100 LeuProThrMetIleValLeuPheLeuLeuValValIleVal 116
|||||
1145 TTGCTTGGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1194
116 ValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArg 133
|||||
1195 GGTGTGCAGATCCGAAGAGCTCCAGACTCTCAAAAAGGGCCCCCGCC 1244
133 LAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
|||||
1245 AGGATCCAGCGCCATAGTGAAGAGCGGGGCTGAAGAGTCCCTGACT 1294
150 ProThrGlnAsnArgLysLysTrpIleTyrcysAsnGlyHisGlyI 166
|||||
1295 CCCACCCAGAACCGGGAATGATCTACTACCGCAACGCGCTCATGTAT 1344
166 eAspIleLeuLysLeuValAlaIleValGlySerGlnTrpLysAspI 183
|||||
1345 TGACATCTTGAACCTGTAGCACCCAGTGGGGAAGCCAGTGAAGAGACA 1394
183 IeTyrcIlnPheLeuCysAsnAlaSerGluArgLysValAlaAlaPheSer 199
|||||
1395 TCTATCAGTTCTTTTCAACGCGCAGTGAAGAGGAGTGGCGCTTCTCC 1444
200 AsnGlyTyrcThrAlaAspHisGluArgAlaTyrcAlaAlaLeuGlnHisTr 216
|||||
1445 AATGGATACCTGCAATCATYGAACGGGCTACGGGCTCTGCAGCACTG 1494
216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
|||||
1495 GACCAATCCGTGGCGCTGAGGCGACGCTTGCCAGCTCATTAGCCCTTGC 1544
```

```

233 rGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
|||||
1545 GCCAGCACCGACCAATGATGTGTGAGAAATGCTGTGGGCTGATGAGAA 1594
250 AspThrThrGlnLeuGluThrHisAspLysLeuAlaLeuProMetSerPro 266
|||||
1595 GACACCAACGAGTTGGAAACAGAACAACTGCTCTCCCACTGAGCCCCAG 1644
266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGlnAsn 283
|||||
1645 TCCGCTGAGCCCGACGCCCATCCCACTGCTTACAGTGAAACTTGAGAAAT 1694
283 eAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
|||||
1695 CCACCTCTCCTGACAGTGAAGCCCTCACCGCTGGAAGAAACAGTCTTC 1744
300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerGlu 316
|||||
1745 TTGCTGAGCAGATCAGAGCCCTTCTGCTTGGAGCTCCACATCCAGTGG 1794
316 YSerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
|||||
1795 CTCTTACGACTGACGACGAAACGGCTCTTTATTACCAAAAGAAAGAGAG 1844
333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
|||||
1845 ACACAGTGTGGCGAGGTGCGCTGAGCCCTGACTGACTGACAGCCCATC 1894
350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleG 366
|||||
1895 TTGTGATGACATGCTGCAATATCTGAACCCCGAGAGACTGCGGGGTATGA 1944
366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
|||||
1945 AGAGATTCCCGACGCTGAGGACAACTGGACCGGCTCTTGCATATCATTTG 1994
383 LysValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrcSer 399
|||||
1995 GGGTCAAGAGCCAAAGAACGACGACAGCCCTTGTGAGCTCTGTGTACAG 2044
400 HisLeuProAspLeuLeu 405
|||||
2045 CATCTTCTGACTTATTG 2062

seq_name: /cgn2_6/plodata/2/1na/6B_COMB.seq:US-09-042-785A-24
seq_documentation_block:
; Sequence 24, Application US/09042785A
; Patent No. 6194151
; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J
; TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,785A
; FILING DATE: 17-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/938,896
; FILING DATE: 26-SEP-1997
; ATTORNEY/AGENT INFORMATION:
```

NAME: Mandragouras, Amy E
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MET-001CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1815 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1815
US-09-042-785A-24

alignment_scores:

Quality: 1527.00 Length: 294
Ratio: 5.194 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-733-956-2 x US-09-042-785A-24

Align seg 1/1 to: US-09-042-785A-24 from: 1 to: 1815

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
|||||
751 ATGAACTCAACAGAAATCCACTTTCGCTCTGTAGACCAAGGATCT 800
17 UserSerIleGlnGluThrValProAspAsnThrSerSerLarg 34
|||||
801 GAGTAGATCATCAGAGAGGAGCAGTCCCTGACACACAGCTCAGCAGG 850
34 LysGluAspValAsnLysThrLeuProAsnLysGlnValAlaAsnHis 50
|||||
851 GGAAGAGAGACGTGAACAGACCTCCCAACCTTCAGTAGTAAACCCAC 900
51 GlnGlnLysProHisLysArgHisIleLeuLysLeuLeuProSerMetG 67
|||||
901 CAGCAAGGCCCCACACAGACATCTCTGAAGCTGCTGCCGTCATGGA 950
67 ValaThrGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
951 GGCACACTGGGGGCGAAGTCCACAGCCCATCAAGGGCCCAAGAGG 1000
84 LysIleProAlaGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
|||||
1001 GACATCTCTAGACAGAACCTACACAGCATTTTGACATCAATGAGACATT 1050
101 ProTrpMetIleValLeuPheLeuLeuValLeuValaIleVala 117
|||||
1051 CCCCGCATGTGATTGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
117 LysSerIleArgLysSerSerArgThrLeuLysGlyProArgGlna 134
|||||
1101 GTGAGATATCCGGAAAGCTCGAGAGCTCTCAAAAGGGGGCCCGCAGG 1150
134 sppsSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1151 ATCCCATGTCATTTGTGAAAGAGCAGGCGCTGAAGAAATCCATGACTCA 1200
151 ThrGlnAsnArgLysLysTrpIleTyTyrCysAsnGlyHisGlyLeas 167
|||||
1201 ACCCAGAAACCGGAGAAATGATCTACTACTGCAATGCGCATGTATCGA 1250
167 PileLeuLysLeuValaIleAlaGlnValGlySerGlnTrpLysAspIleT 184
|||||
1251 TATCTGAAGCTGTAGCAGCCCAAGTGGAGGACAGTGTGAAAGATATCT 1300

seq_name: /cgn2_6/plodata/2/ina/5B_comb.seq:us-09-042-785A-22

seq_documentation_block:

Sequence 22, Application US/09042785A

Patent No. 6194151

GENERAL INFORMATION:

APPLICANT: Busfield, Samantha J

TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,785A

FILING DATE: 17-MAR-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/938,896

FILING DATE: 26-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MET-001CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)742-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 2638 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 510..2327

US-09-042-785A-22

alignment_scores:
 Quality: 1527.00 Length: 294
 Ratio: 5.194 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-733-956-2 x US-09-042-785A-22

Align seg 1/1 to: US-09-042-785A-22 from: 1 to: 2638

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLeu 17
  |||||
1260 ATGAACCTCAACAGAAATCCAACTCTTCTGCTCTGTAGACCAAGGTA 1309
  |||||
17 uSerSerIleGlnGluGlyThrValProAsnThrSerSerAlaArgG 34
  |||||
1310 GAGTAGCATCCAGGAAGGAGAGTCCCTGACAAACAGCAAGCTAGCAAGG 1359
  |||||
34 1YLSGLuAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
  |||||
1360 GGAAGGAACAGTGAACAGACCTCCCAACCTTAGAGTACCAACAC 1409
  |||||
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMet 67
  |||||
1410 CAGCAAGGCCCCACCACAGACATCCCTGAAGCTGCTCCGTCATGGA 1459
  |||||
67 uAlaThrGlnGlyGluLysSerSerThrProIleLysGlyProLysArg 84
  |||||
1460 GGGCCATGGGGGGGAGAAAGTCCAGCACGCCCATCAAGGGCCCCAGAG 1509
  |||||
84 1YH1SProArgLnsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
  |||||
1510 GACATCCCTAGACAGAACTCACAACAGCATTTGACATCAATGACATTG 1559
  |||||
101 ProThrMetIleValLeuPheLeuLeuValLeuValIleValVal 117
  |||||
1560 CCTGTATATGTGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
  |||||
117 1GYSerIleArgLysSerSerArgThrIleuLysLysGlyProArgGln 134
  |||||
1610 GTGCAGTATCCGGAAAGCTCCAGGAGCTGTAAAGGGGGCCCCGCGCAG 1659
  |||||
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
  |||||
1660 ATCCCACTGCCATTGTGAAAGGCGGCTGAAGAAATCCATGACTCCA 1709
  |||||
151 ThrGlnAsnArgGluLysThrPrlIleTyrrYrcysAsnGlyHisGlyTleAs 167
  |||||
1710 ACCCGAAGCCGGGAGAAATGATCTACTACTCAATGGCCATGATATCGA 1759
  |||||
167 PileLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
  |||||
1760 TATCCGGAAGCTTGTAGCAAGCCCAAGTGGGAAGCCGATGGAAGAATCT 1809
  |||||
184 yGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
  |||||
1810 ATCAGTTTCTTTCGAATGCCAGTAGAGAGGAGTTGCTGCTTCTTCAT 1859
  |||||
201 GlyTyrrThrAlaAspHisGluArgAlaTyrrAlaAlaLeuGlnHisTrpH 217
  |||||
1860 GGGTACACAGCGAGCACGAGCGGGGCTAGCGAGCTCTCAGACACTGGAC 1909
  |||||
217 rIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
  |||||
1910 CATCCGGGGGGCCGAGGCCAGCTCGCCAGCAATTAATGCCCGCTCGGCC 1959
  |||||
234 1NH1sArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
  |||||
1960 ACCACCGGGAACAGATGTTGTGAGAAAGATTCGTGGGCTGATGGAAGAC 2009
  |||||

```

```

251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
  |||||
2010 ACCACCGAGCTGGAAACTGACAACTAGCTCTCCGATGAGCCCGACGCC 2059
  |||||
267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGlnAsnSerA 284
  |||||
2060 GCTTAGCCCGAGCCCATCCCAAGCCCAAGCGAAACTTGAGAAATTCGG 2109
  |||||
284 1aLeuLeuThrValGluProSerProGlnAsp 294
  |||||
2110 CTCTCTGACGGTGAGAGCTTCCCGACAGAGAT 2141
  |||||

```

seq_name: /cgn2_6/prodata/2/ina/6B_COMB_seq:US-09-042-785A-3

seq_documentation_block:

: Sequence 3, Application US/09042785A

: Patent No. 6194151

: GENERAL INFORMATION:

: APPLICANT: Busfield, Samantha J

: TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY

: NUMBER OF SEQUENCES: 31

: CORRESPONDENCE ADDRESS:

: ADDRESS: LAHIVE & COCKFIELD, LLP

: STREET: 28 State Street

: CITY: Boston

: STATE: Massachusetts

: COUNTRY: USA

: ZIP: 02109

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: PatentIn Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/09/042,785A

: FILING DATE: 17-MAR-1998

: PRIORITY APPLICATION DATA:

: APPLICATION NUMBER: US 08/938,896

: FILING DATE: 26-SEP-1997

: ATTORNEY/AGENT INFORMATION:

: NAME: Mandragouras, Amy E

: REGISTRATION NUMBER: 36,207

: REFERENCE/DOCKET NUMBER: MET-001CP

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (617)227-7400

: TELEFAX: (617)742-4214

: INFORMATION FOR SEQ ID NO: 3:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 2612 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: single

: TOPOLOGY: linear

: MOLECULE TYPE: CDNA

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 190..951

: US-09-042-785A-3

alignment_scores:

Quality: 1269.00 Length: 259
 Ratio: 5.138 Gaps: 1
 Percent Similarity: 95.367 Percent Identity: 94.981

alignment_block:

US-09-733-956-2 x US-09-042-785A-3

Align seg 1/1 to: US-09-042-785A-3 from: 1 to: 2612

```

11 SerValArgProLysValLeuSerSerIleGlnGlyThrValProAs 27
  |||||
930 TCCCAAGACAAAGGTAAGTACTGAGTACATCCAGGAAGGAGAGTCCCTGA 979
  |||||

```



```

ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,111
FILING DATE: 06-NOV-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02536
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 6..6554
US-08-746-111-4

Alignment scores:
Quality: 115.50 Length: 488
Ratio: 0.516 Gaps: 20
Percent Similarity: 45.902 Percent Identity: 19.877

Alignment block:
US-09-733-956-2 x US-08-746-111-4 ..

Align seg 1/1 to: US-08-746-111-4 from: 1 to: 6585

5 GluSerAsnSerSerIAsaValArgProLysValLeuSerSerIlegI 21
:::|||||
2322 GAGCTCAACTCTCA.....CGAATCTTAGTAATAATCAT 2356
21 ngluglythrvalproasp.....asnhrSerSerIArgGlyLysG 36
:::|||||
2357 CAATATTAACCTTCAAAGACTTCACAAAGAACACTTCGCTCAGGAGCA 2406
36 lnaSprAlaSnlyThrLeuProAsnLeu..... 45
CCGTGGCTGGTACCTCTTAGAAGACTTCAGTGGCTTAGATGAGAACTTC 2456
46 .....GlnValAlaSnHISGlnGlnGlyProHISHisArgHIS.. 58
:::|||||
2457 GTCCTCAACTCTTCTACAGAACATCGTCCAGCTCAATCATGAAAAATGA 2506
59 .....IleLeuLysLeuLeuProSerH 66
TATGAAAAATCCACAGTCAAAACATCAACAATGATATACCTACTCTCTTG 2556
66 eGcluaIaThrGlnGlyGlnLysSerSerThrProIleLysGlyProLys 82
:::|||||
2557 GTCCAAAGAGATCTGGGAATCGAGAACACAAATATAACTAAACATCAAG 2606
83 ArgGlyHisProArgGlnAlaSnLeuHisLysHisPhe..... 94
:::|||||
2607 ACAGGAAGACCCCACTGATGAAGACACACAGGTTCTCTGATGAAGACGCC 2656

```

```

95 .....AspIleasnGluHisLeuProTrpMetIleValLeuPheLeu 109
2657 AGCTGGTAAACCTGGAGGCGAT..... 2678
109 euLeuValLeuValIleValValCysSerIleArgIysSerSerArg 125
2679 .....TCAACCCCAAGAAATGCTAT 2699
126 ThrLeuLysLysGlyProArgGlnAspProSerAlaIleValGluLysAl 142
2700 TCTGGAAATAGCTGAGGAGGACATTCCTAGCGAGTGTATACC..... 2744
142 agLysLeuLysSerMetThrProThrGlnAsnArgGluLysTrp... 157
2745 ....TTAAAGCAAAAGATCATCTCCAAATTTCTGAATAGACGATGCGCTG 2790
158 .....IleTyTrpCysAsnGlyHisGly 165
2791 TGGCTTCTGAAAAGGCTAGTTATGAATATATACAGCAATGGGTGAGAC 2840
166 IleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAs 182
2841 ACAGATGTGGATTAAGCTGACC..... 2861
182 PileTyTrpGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheS 199
2862 .....ACAGTCTCCAAATCAGAAATATACAGTAC 2892
199 eArgsnGlyTrpThrAlaAspHisGluArgAlaValAlaLeuGlnHis 215
2893 CTCGGGAGAGACACCTCTCACACAAC..... 2921
216 TrpThrIleArgGlyProGlu.....AlaSerLeuAlaG 227
2922 ...ACAACAAGAAAGCCAGTACCTCCCAACATTTCTCGAGTGTGACA 2968
227 nLeuIleSerAlaLeuArgGlnHisArgArgAsnAspValValGluLysI 244
2969 TTAATCTCACATGTAAAGACAGAGAGAGAAACAGTGTTCACAGAAA 3018
244 leArgGlyLeuMetGluAspThrThrGlnLeuThrAspLysLeuAla 260
3019 GACAGTATTATTCACAGACACGAGAGAGAGAGAAATAAGAGCTTGCA 3068
261 Leu.....PrometSerProSerProLeuSerPro.....Se 271
3069 CTACACAGTCTCTATCTCCAGGGCTTGACCTTTGAGAGACATTA 3118
271 rProIleProSerProAsnAlaLysLeuGlnAsnSerAlaLeuLeu... 286
3119 CCATTCCCATTTCCAGACAGAGACTACTTAATCACTCACTGTTCTCTCC 3168
287 .....ThrValGluProSerProGlnAspLysAsnLys 297
3169 ACAAGTCAATGAACAGCTCTTCTCCAGACTGAACCAAGACTCTCTCC 3218
298 GlyPhePheValAspGluSerGluPro.....LeuLeuArgCysAs 311
3219 TCAATGATGACGAGCAGTCACTCTCTGACTATATACAGTACTCGAAAA 3268
311 pSerThrSerSerGlySerSerAlaLeuSerArgAsnGlySerPheLeu 328
3269 TGCAGCTGAGCAGATGAGCTCTTCTTAAGTCTTTATCACTAGTGCCTCC 3318
328 hArgSGLuLysLysAspThrValLeuArgGln..... 338
3319 CAGAGGAACACTCTCCACATTTCTGCGCCCAAGATCTGATCAAAACAC 3368
339 .....ValArgLeuAspProCysAspLeuGlnProI 349
3369 TCTACACAGATCTCTAGTACAGATCTCTCCGCGCAGAGCTCAGCCAGG 3418

```

```

349 ePheAsp...AspMetLeuHisPheLeuAsnProGluGluLeuArgValI 365
3419 GCTTGATATAGACCTTAAGTCATGACTTTTACCTGTGATGACATGTCTAA 3468
365 leGluGluIleProGlnAlaGluAspLysLeuAsp..... 376
3469 CATCTTTCTTTCCAGACCAAGTCAAAAGTCATCTTCTCTCAGATGAT 3518
377 .....ArgLeuPheGluIleIleGlyVal 384
3519 GACCAAGCAATCCCTCTCCAGACTTAAGCTCTTACCATC..... 3560
384 LysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTySerHisL 401
3561 ....TCTCCAGAAATGAGACAGACAAT.....ATTACCCAGACC 3597
401 euProAspLeuLeu 405
3598 TGGATCAGTTGCTC 3611

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-041-886-1
seq_documentation_block:
: Sequence 1, Application US/09041886
: Patent No. 6235872
: GENERAL INFORMATION:
: APPLICANT: Bredesen, Dale E.
: APPLICANT: Rabizadeh, Sharroz
: TITLE OF INVENTION: Proapoptotic Peptides, Dependence
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESS: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: United States
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/041,886
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 2626
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-6949
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3386 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 114..1395
: US-09-041-886-1

alignment_scores:
: Quality: 113.00 Length: 300
: Ratio: 0.837 Gaps: 12
: Percent Similarity: 45.000 Percent Identity: 22.333

alignment_block:

```

US-09-733-956-2 x US-09-041-886-1 ..

Align seg 1/1 to: US-09-041-886-1 from: 1 to: 3386

```

62 LeuLeuProSerMetGluAlaThrGlyGlyIuLysSerSerThrProI 78
   ||| ||||| ||| ||||| |||||
910 TTGTGGCTACATAGCTTCAAGAGTGGAAAGACTGTCAACACACACAG 959
    eLysGlyPro.....LysArgGlyHisProArgGln 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
960 CAAGAGACCAACACCGCCAGTACACAGAGCCGCCACAGAGGAG. 1008
    snLeuHisLysHisPheAspIleasnGluHisLeuProTrpMetIleVal 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1009 .....AAAACTCCACAGCCAGATGGCATCTCCGTG..... 1041
    LeuPheLeuLeuValIleValIleValIleValIleValIleValIleVal 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1042 .....ACAGCCACAGCCCTGCAT 1058
    sSerSerArgThrLeuLysGlyProArgGlnAspSerAlaIleVal 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1059 GACACAGACCCACACAGCAGACAGCTCGGGCCAGGCCCTCAAGAGTGA 1107
    alGluLysAlaGlyLeuLysSerMetThrProThrGlnAsnArgGlu 155
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1108 C.....GGAGGCTCTACAGAGCTGGCCCGCCAGCCACAGGGAGAG 1151
    LysTrpIleTyTyrCysasnGlyHisGlyIleAspIleLeuLysLeuVal 172
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1152 .....GTGGAGAGCTCT 1165
    lAlaAlaGluValGlySerGlnTrpLysAspIleTyTyrGlnPheLeuVal 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1166 CAAGGCTCTGGGGGACACCTGGCGCAGCTG..... 1199
    snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyTyrThrAlaAsp 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1200 .....GGGGGAGAGTGGGCTACACAGCCGAG 1226
    HisGluArgAlaTyTyr.....AlaAlaLeuGlu 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1227 CACATAGACTCTTACCATGAGCGCTGCCCTGGCGCTGCTGCTGCTGC 1276
    nHisTrpThrIleArgGlyProGluAlaSerIleAlaGlnLeuIleSerAla 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1277 AAGCTGGGCCACCCAG...GACAGCGCCACACTGGAGCCCTCTCTGGCG 1333
    lAlaLeuArgGlnHisArgAlaGlnAsnAspValValGluLysIleArgGlyLeu 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1324 CCCTGGCGCCGATCCAGCAGCGACCTCGTGAG.....AGTCTG 1364
    MetClnAspThrThrGlnLeuGluThrAspLysLeuAlaLeuPrometSe 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1365 TCGAGTAGTCCACTGCCACATCC.....CCGCTGCTG 1396
    rProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG 281
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1397 AGCCCAACCGGGAGCCCGCCGCCGCCCATCTCCGACACAGCGATGCT 1446
    lAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLys 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1447 CCAAGCAACCCCTGTGGAGCCCGACCCCGCTTGGGGGGGGCCCG 1496
    GlyPhePheValAspGluSerGluProLeuLeuArgCysAspSerThrSe 314
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1497 CTGGCAGAACTGAGCTCTGTGGCAGAGCTCAGAGTCCAGGCCCA 1546
    rSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGlu 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1547 ACACAGCCCTGTGACAGCAGCCCGCTGTGGCCCTTC...ACTTCTGACC 1593
    yLysAspThrValLeuArgGlnValArgLeu.....AspProCys 344

```

```

1594 AACACTTCTGTCCAGAGAGAGAGAGAGTGGCCCTGCTGCCCTCCCAACCCCTGC 1643
   ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
seq_name: /cgn2_6/prodata/2/ina/6A_COMB.seq:US-08-602-791-1

```

```

seq_documentation_block:
; Sequence 1, Application US/08602791
; Patent No. 6074836
; GENERAL INFORMATION:
; APPLICANT: BORDIGNON, Claudio
; APPLICANT: MAVILLO, Fulvio
; TITLE OF INVENTION: METHOD OF MARKING EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIIDO, MARMELESTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N.W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,791
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM93/A000587
; FILING DATE: 01-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/06723
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, Sharon N.
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: P1614-4007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: one-of(943)
; OTHER INFORMATION: /function= "pvuII cleavage site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: one-of(1512)
; OTHER INFORMATION: /function= "SctI cleavage site"
US-08-602-791-1

```

```

alignment_scores:
Quality: 112.00 Length: 245
Ratio: 1.008 Gaps: 10
Percent Similarity: 45.306 Percent Identity: 22.449

```

```

alignment_block:
US-09-733-956-2 x US-08-602-791-1 ..

```

```

Align seg 1/1 to: US-08-602-791-1 from: 1 to: 1600

```

```

62 LeuLeuProSerMetGluAlaThrGlyGlyIuLysSerSerThrProI 78
   ||| ||||| ||| ||||| |||||

```

```

910 TTGTGGCCTACATAGCCTTCAGAGGTGCAACAGCTGCAAGCAAGCAAG 959
78 elysglyPro.....LysArgGlyHisProArgGlnA 89
   ||| |||
   ||| ||| ||| |||
960 CAAGAGGCCACAGCGCGGCGAGTGAACAGAGCCGCCACCAAGAGGAG. 1008
89 snLeuHisLysHisPheAspIleAsnGlnHisLeuProTrpMetIleVal 105
   ||| |||
   ||| ||| ||| |||
1009 .....AAAACTCCACAGCCAGTGGCAGTCCTCGTGG..... 1041
106 LeuPheLeuLeuLeuValIleValIleValIleValCysSerIleArgly 122
   ||| |||
   ||| |||
1042 .....ACAGCCAGAGCTGCAT 1058
122 sSerSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleVal 139
   ||| |||
   ||| ||| ||| |||
1059 GACCACAGAGCCGCCACGACGAGACGCTCGCGCCAGCCCTCAAGGTTGA 1107
139 aIGluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGlu 155
   ||| |||
   ||| ||| ||| |||
1108 C.....GGAGCCTCTACAGCAGCCTGCCCCAGCAAGCGGAGGAG 1151
156 LysTrpIleTyTrpCysAsnGlyHisGlyIleAspIleLeuLysLeuVal 172
   ||| |||
   ||| ||| ||| |||
1152 .....GTGAGAGGCTTCT 1165
172 LAlaAlaGlnValGlySerGlnTrpLysAspIleTyTrpGlnPheLeuCysA 189
   ||| |||
   ||| ||| ||| |||
1166 CACGGGCTCTGGCGGGGACACCTGGCGGACCTG..... 1199
189 snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyTrpAlaAsp 205
   ||| |||
   ||| ||| ||| |||
1200 .....GCGGGCGAGCTGGGCTGCCAGCCCGAG 1226
206 HisGluArgAlaTyr.....AlaAlaLeuGlu 214
   ||| |||
   ||| ||| ||| |||
1227 CACATGAGCTCTTTACCCATGAGGCTGCCCGCTCGCGCCTGCTTGC 1276
214 nhIstPrpThrIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerA 231
   ||| |||
   ||| ||| ||| |||
1277 AACCTGGGCGCACCCAG...GACAGCGCACACTGGACGCCCTCCTGGCCG 1233
231 LalaLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
   ||| |||
   ||| ||| ||| |||
1324 CCTGGGCGCATCGACGCGGACCTCGTGAG.....AGCTCG 1364
248 MetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSe 264
   ||| |||
   ||| ||| ||| |||
1365 TGCAGTGAAGCCACATGCCACATCC.....CCGGGTG 1396
264 rProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG 281
   ||| |||
   ||| ||| ||| |||
1397 AGCCCAACCGGGAGCCCGCCGCCGCCACATCCGACCAACCGATGCT 1446
281 LuAsnSerAlaLeuLeuThrValGluProSerPro 292
   ||| |||
   ||| ||| ||| |||
1447 CCAGCCAAACCCCTGTGGAGCCGACCCGCCACCCCT 1481
seq_name: /sgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-985-950-5
seq_documentation_block:
; Sequence 5, Application US/08985950
; Patent No. 6140076
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,950
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,279
; FILING DATE: 21-MARCH-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,181
; FILING DATE: 16-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,252
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0670K
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1204
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 155..1015
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1247
; OTHER INFORMATION: /note="nucleotide 1247 designated
; NAME/KEY: mat_peptide
; LOCATION: 218..1015
; US-08-985-950-5

alignment_scores:
      Quality: 107.50      Length: 338
      Ratio: 0.773        Gaps: 17
      Percent Similarity: 41.124      Percent Identity: 20.414

alignment_block:
align seg 1/1 to: US-08-985-950-5 from: 1 to: 1279
27 AspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuPr 43
   ||| |||
   ||| ||| ||| |||
341 GAGAGTAGATGCACATACATATGTAAGAGGAGGAGGAGGAGGAGGAGG 390
43 OAsnLeu.....GlnValAlaAsnHisGlnG 52
   ||| |||
   ||| ||| ||| |||
391 ATCTGAGTCAGAGCCAGATTCGCGATTCGACTGACTGAGTGAAGGAAATG 440
52 InGlyProHis.....His 56
   ||| |||
   ||| ||| ||| |||
441 CCGGCGCTTATCGCTCATCTATATTAAGCCCTTAATAGTGTGAGCGAG 490
57 ArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGlyGlyGlu 73
   ||| |||
   ||| ||| ||| |||
491 AGTGACTACTGTGAGCGTGGTGA...GAAACCTGTGAGAGCCCGGGA 537
73 sSerSerThrProIleLysGlyProLysArgGly.....HisProA 87

```

[illegible]

```

1  APPLICANT: Adema, Gosse Jan
2  TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes
3  NUMBER OF SEQUENCES: 22
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: DNAX Research Institute
6  STREET: 901 California Avenue
7  CITY: Palo Alto
8  STATE: California
9  COUNTRY: USA
10 ZIP: 94304-1104
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.30
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/985,950
20 FILING DATE: 05-DEC-1997
21 CLASSIFICATION: 435
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 60/041,279
25 FILING DATE: 21-MARCH-1997
26
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 60/033,181
29 FILING DATE: 16-DEC-1996
30
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 60/032,252
33 FILING DATE: 06-DEC-1996
34
35 ATTORNEY/AGENT INFORMATION:
36 NAME: Ching, Edwin P.
37 REGISTRATION NUMBER: 34,090
38 REFERENCE/DOCKET NUMBER: DX0670K
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE: (650)852-9196
41
42 TELEFAX: (650)496-1204
43
44 INFORMATION FOR SEQ ID NO: 7:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 1728 base pairs
47 TYPE: nucleic acid
48 STRANDEDNESS: single
49 TOPOLOGY: linear
50
51 MOLECULE TYPE: cDNA
52
53 FEATURE:
54 NAME/KEY: CDS
55 LOCATION: 69..929
56
57 FEATURE:
58 NAME/KEY: mat_peptide
59 LOCATION: 132..929
60
61 US-08-985-950-7
62
63 alignment_scores:
64
65 Quality: 107.50 Length: 338
66 Ratio: 0.773 Gaps: 17
67 Percent Similarity: 41.124 Percent Identity: 20.414
68
69 alignment block:
70 US-09-733-956-2 x US-08-985-950-7 ..
71
72 Align seg 1/1 to: US-08-985-950-7 from: 1 to: 1728
73
74 27 ASPAAnThrSerSerIaArgGlyLysGluASpValAsnLysThrLeupr 43
75 ::::: |||::: ::::: |||||::: ::::: ||
76 255 GAGAGTAAAGTCCACATACATATGATCTACTGAAGATGTGTCTCAACTGATCC 304
77
78 43 oASnIeu.....GlnValValAsnHisGlnG 52
79 ::::: |||::: ::::: |||:::
80 305 ATCTGATGTCAGAGGCCAGATTCCGACATTGACTCAGTAAAGTGAAGAAATG 354
81
82 52 InGlypHnHs.....HIs 56
83 |||||::: :::::
84 355 CCGGGCCCTTAATGCTGCATCTATTATTAAGCCCCCTAAATGCTGTGAGCAG 404

```

```

57  ArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGlyGlyGly 73
    |||:|||||
405  AGTACTACTGAGCTGCTGAGTGA...GAAACCTCTGAGGCCGCA 451
    |||:|||||
73  sSerSerThrProIleGlyGlyProLysArgGly.....HisPro 87
    |||:|||||
452  CTCCCGGACACAGACCCGGCTCTAGCTGAGCCACGACGAGCCGCT 501
    |||:|||||
87  rglInuLeuHisLysHisPheAspIleasnGluHisLeuPro..... 101
    |||:|||||
502  CGGACAAACAGTAC.....AATGACATGACCTGCTGCC 536
    |||:|||||
102  .....TpmIleValle 106
    |||:|||||
537  CAAGGCGTGAACCTGACATCTGTATTTCTCATCGGGCTCTCAGTGT 586
    |||:|||||
106  uPheLeuLeuValLeuValValIleValValCysSerIleArgLys 123
    |||:|||||
587  CTTCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 636
    |||:|||||
123  eSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleVal 139
    |||:|||||
637  ATCAG.....ATAAGCAGGGGCCCCCAGAGCAAGCAGC..... 671
    |||:|||||
140  GluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGly 156
    |||:|||||
156  strPleTyTYTCysasnGlyHisGlyIleAspIleLeuLysLeuVal 173
    |||:|||||
704  G.....GCTGTGATGTTCTTGAAGAGCAGCAG 729
    |||:|||||
173  laAlaGlnVal.....GlySerGlnTrpLysAspIleTyrglnPhe 186
    |||:|||||
730  CAGACAAGGCCACATCAATGACTTCTCTGAGAAGC..... 767
    |||:|||||
187  LeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyTh 203
    |||:|||||
768  .....AGAGAGCGGACACCTCGCCCTGCTCGCAGG..... 800
    |||:|||||
203  rAlaAspHisGluArgAlaTyAlaAlaLeuGlnHisTrpThrIleArg 220
    |||:|||||
801  .AGTTCCAGAGGTAGCTATGCTCAGCTGACACACAGTGGCCCTC... 845
    |||:|||||
220  lProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArg 236
    |||:|||||
845  ..... 845
    |||:|||||
237  ArgAsnAspValValGluLysIleArgGlyLeuMetGluAspThrThr 253
    |||:|||||
845  ..... 845
    |||:|||||
253  nLeuGluThrAspLysLeuAlaLeuProMetSerPro.SerProLeuSer 269
    |||:|||||
846  .....ACACAGAGGACACCCGGGTGTGTCCACACAGTCCACAAGC 888
    |||:|||||
270  ProSerProIleProSerProAsnAlaLysLeuGluAsnSerAlaLeu 286
    |||:|||||
889  CCATGGCCGAGTCATCA..... 906
    |||:|||||
286  uThrValGluProSerProLysAsnLysGlyPhePheValAsp 303
    |||:|||||
907  .CGTATGACGCCCTTGCCA.....G 925
    |||:|||||
303  lSerGluProLeu.....LeuArgCysAspSerThr 313
    |||:|||||
926  ACACGTGACCCATACCCACCTGGCTCTGTGACCTGAGGAGTGAAGTCA 975
    |||:|||||
314  SerSerGlySer 317
    |||:|||||
976  TCTAGGAAAAGC 987
    |||:|||||

```

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-961-083-55

```

seq_documentation_block:
; Sequence 55, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Chol et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-55

alignment_scores:
    quality: 105.00    length: 534
    ratio: 0.482      gaps: 23
    percent similarity: 40.824    percent identity: 17.978

alignment_block:
US-09-733-956-2 x US-08-961-083-55 ..

Align seg 1/1 to: US-08-961-083-55 from: 1 to: 2389

5  GluSerAsnSerSerAlaSerValArgProLysValLeuSerSerIleG 21
   |||:|||||
41  GAAATATATGCTGTTCTCTATATAGATGAAACACAGCGCAAAAC 90
   |||:|||||
21  nGluGlyThrValProAspAsnThrSerSerAlaArgGly..... 34
   |||:|||||
91  GGAGAAATTTCATCTGATGAGGTAGCAAGCGTGAAGAAATCAATGCTG 140
   |||:|||||
34  ..... 34
   |||:|||||
141  AGCAATCGTCATCAAGATTAACAGACCAAGCTATGTCACTTCACATGGC 190
   |||:|||||
34  ..... 34
   |||:|||||
191  GACCACATCATATTATCAATGCTAAGGTTCTTATGAGCGTATCATCAG 240
   |||:|||||
35  .....LysGluAspValAsnLysThrLeuProAsnLeuGln 47
   |||:|||||
241  TGAAGAATTTCATCATGAAGATCCAAACTATATAGCTTAAAGATGAGGATA 290
   |||:|||||

```

```
47 aValAsnHisGlnGlyProHisHisArgHisIleLeuLysLeu... 62
   |||||
291 TTGTTAATGAGGTCAAGGT.....GGATGTTATCAAGTTGAT 331
63 .....LeuProSerMetCysAlaThrGlyGlyGly 73
332 GGAATACTATGTTTACCTTAAGATGTCGCCACCGGATACGCCG 381
73 sSerSerThrProIleLysGlyProLysArgGlyHisProArgGlnAsnL 90
   |||||
382 TACAAAGAGAGAAATCAATGCACAAACAGACCATAGTCA..... 424
90 eunHisLysHis.PheAspIleAsnGlnHis.....LeuProTrp.. 102
   |||||
425 ..CATGCTGAAGTGAAGTCCAGAACATGAGTGTGCTGCTGCTGGCA 472
103 .....MetIleValLeuPheLeuLeuL 110
   |||||
473 CGTTCCGACAGACGCTTACTACAGATGATGTTATGTTTAAATGCTTC 522
110 euVal.....LeuValIle 115
   |||||
523 TGAATGATGAGAGATCTGCTGATCTTATGCTTCATGAGATC 572
116 ValValCysSerIleArgLysSerSerArgThrLeuLysGlyProAr 132
   |||||
573 ATTACCACTTACATTCCTAAGATGATTCACGATGACGAGTGGCTGCT 622
132 gGlnAspProSerAlaIleValGlu.LysAlaGlyLeuLysLysSerMet 148
   |||||
623 GCAGAACCTTCCTCTGCTGCTGCAGAAATCTCAAAATTCAGAACCTA 672
149 ThrProThrGlnAsnArg.....GluLys.TripLe..... 158
   |||||
673 TCGCGACAAATAGCATTAACCTTCAAGACAACTGGGTACTCTTCG 722
159 .....TyrTyrCysAsnGlyHisGlyIle 166
   |||||
723 TAAGCAATCCAGAACTACAATACTAACAACCAACCAACAGCAACACT 772
167 AspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
   |||||
773 AACGTCACACCAAGTCAAGTAATGACATGATGCTCTTGAACAGCT 822
183 eTyrGlnPheLeuCysAsnAlaSerGluArgGluVal..... 195
   |||||
823 CTACAAACTG.....CCTTGAAGTCAACGACATGTAAGATCTGATGCC 866
196 .....AlaAlaPheSerAsnGlyTyrThr 203
   |||||
867 TTGCTTTGATCCAGACAATCAACAGTGCACAGCTAGAGGTGGCA 916
204 AlaAspHisGlnArgAlaTyrAlaAlaLeuGlnHisTrpHisArgGly 220
   |||||
917 GTGCACACGAGATCATACCACTTACCTTACCTCAAAATGCTGTA 966
220 yProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArg 237
   |||||
967 ATTGGAAGAACGAATCGCTGATATTATTCCTT.....CGTTATGCTT 1010
237 rGlnAsnArgValIleGluLysIleArgGlyLeuMetGluAspThrThrGln 253
   |||||
1011 CAACACCATGGGTACAGATTCAAG..... 1036
254 LeuGlnThrAspLysLeuAlaLeuPheMetSerProSerProLeu.... 268
   |||||
1037 .....CCAGAACCAACCAAGTCCACCAACGAC 1062
269 .....SerProSerProIleProSerProAsnAlaLysLeuGln 282
   |||||
1063 TCCGGAACCTAGTCCAGCCCGCAACTGCACCAATCTTAAATAGACT 1112
282 snSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGly 298
```

```
1113 CAAATCTTCTTGG...GTAGTAGAGTGTGTAACGAAAGTGGGGAAGCA 1159
   |||||
299 PhePheValAspGlu..... 303
1160 TATGTATTCCGAAGAAAGGCGATCTCTGTTATGCTTTGGCAAGATT 1209
304 ....SerGluProLeuLeuArgCysAspSerThrSerGlySerSer 319
   |||||
1210 ACCATCTGAACCTGTAAAAATCTTGAAGACAGTATTCAAAACAGAGA 1259
319 IaleuSer..... 321
1260 GTGTTTACACACTTTTAACTGCTAAAGAAAGAAATGTTGCTGCTGAC 1309
321 ..... 321
1310 CAAGAAATTTTATGATTAAGCATATATCTGTTAACTGAGGCTCATAAAGC 1359
322 .....ArgAsnGlySerPheIleThrLysGlyLysL 332
1360 CTGTGTTGMAAATAAGGTCGTAATCTGATTTTC.....CAAGCTT 1400
332 yAspThrValLeuLeuArgGlnValAlaLeuAspProCysAspLeuGlnPro 348
   |||||
1401 TAGACAAATTTATTAAGACGCTTGAATGATGATGACATTAATGAAGAAA 1450
349 lIlePheAspSerMetLeuHisPheLeu.....AsnProGluGly 361
   |||||
1451 TTGGTAGATGATTTATTTGCGATCTCTACACCAATTAACCATCCAGAGG 1500
361 uLeu...ArgValIleGluGluIleProGlnAlaGluAspLysLeu 375
1501 ACTTGCGAACCAAAATTTCTCAAAATGAGTAACTGAAGACGAAAGTT 1546
seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:us-08-985-916-15
seq_documentation_block:
; Sequence 15, Application us/08985916
; Patent No. 6221636
; GENERAL INFORMATION:
; APPLICANT: ATSUSHI HAYAKAWA, MASARAZU SUGIMOTO, YASUHIKO YOSHIMURA, AND TSUYOSH
; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1735 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; COUNTRY: VA
; ZIP: 22152
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,916
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-325658
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3533 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```

: TOPOLOGY: linear
: MOLECULE TYPE: genomic DNA
: ORIGINAL SOURCE:
: ORGANISM: Brevibacterium lactofermentum
: STRAIN: ATCC 13869
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 321..3077
: US-08-985-916-15

```

```

alignment_scores:
  Quality: 101.00      Length: 370
  Ratio: 0.580         Gaps: 25
  Percent Similarity: 47.027  Percent Identity: 24.054

```

alignment_block:

US-09-733-956-2 x US-08-985-916-15 ..

Align seg 1/1 to: US-08-985-916-15 from: 1 to: 3533

```

54 ProHisHisArgHisIleLeuLeuLeuProSerMetGluAlaThrG1 70
   ||| ||||| : : : : : ||| ||||| ||
1607 CCAGACACCGCCAAACTACCGCGAGCTGTGAAAGCAGAGAGCTTGAGG 1656
   70 y...GlyIuLysSerSerThrProIleLysGlyPro..... 81
   | : : : : : : : : : : | : : : : :
1657 TCGCTGTAAGAGAACTGGCAGCGCTGCTCCGCTGATCCCGACGGTTCA 1706
   82 .....LysArgGlyHis..ProArgGlnAsnIleHisLysHis..Ph 94
   : : : : : : : : : : : : : : : : : : : :
1707 GATGAATACAGCGAGGTACCGACCGGAGCT.....CGGCATCTT 1747
   94 eAspIleAsnGlnHisLeuProTrpMetIleValLeuPheLeuLeuVal 111
   | : : : : : : : : : : : : : : : :
1748 C..... 1748
   111 AlLeuValValIleValValCysSerIleArgLysSerSerArgThrLeu 127
   ||| : : : : : : : : : :
1749 .....CGCACCGCTCGGAGCGCTGT 1769
   128 LysLys...GlyProArgGlnAspProSerAlaIleValGluLysAlaG1 143
   ||||| : : : : : : : : : : |||
1770 AAGAAATTTGGGCCAGGATGGTCCCTCACTGCATCTTCACAGGA.. 1817
   143 yLeuLysLysSerMetThrProThrGlnAsn.....Arg 155
   : : : : : : : : : : : : : : : :
1818 .....TCATCGCTACCGAGTGTGCTCGAGCGCATGTTCCTCAAG 1860
   155 LuLysTrpIleTyTyCysAsnGlyHisGly.....IleAsp 167
   | : : : : : : : : : : : : : : : :
1861 AATTCCGCTCATTTGCAGCCACAGCGGACAAACCACGCGCACCGTCGAT 1910
   168 IleLeuLysLeuVal.....AlaAlaG1 175
   : : : : : : : : : : : : : : : :
1911 GTATCCCACTGTTGCAAAACCATGAAAGATCTCCAGCGCGCGGGAAT 1960
   175 nValGlySerGlnTrpLys...AspIleTyGln.....PheLeuCysA 189
   : : : : : : : : : : : : : : : :
1961 CCGCAGCAAGAACTGGAAATTTATCTTTACCCGAACTACCTCTGAGC 2010
   189 snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyThrAlaAsp 205
   : : : : : : : : : : : : : : : :
2011 GCGACACAGCTCCAGAAATCAATCTC.....GTTACTCCGATTC 2051
   206 HisGluArgAlaTyTrpAlaAlaLeuGlnHisTrpThrIleArgGlyProG1 222
   : : : : : : : : : : : : : : : :
2052 AACCAAGATGGCGGATATTTCTCGCAAACTGGCGCTTTACAGCGGGA 2101
   222 uAlaSerLeuAlaGlnLeuIleSerAla.....LeuArgGln.. 234
   | : : : : : : : : : : : : : : : :
2102 ACTGCGAGCTCGTGCAGTATGCGGATCAGCGGCGGTCAAGCTTCGCTGT 2151

```

```

235 ..HisArgArgAsnAspValValGluLys..IleArgGlyLeuMetGluAs 250
   ||| ||||| : : : : : |||
2152 TCCAGCGCGCTGGTGGCAGCTGGCGCGGTGGCGGAGCTT.....CC 2195
   250 pThrThrGlnLeuGlnIuThrAspLysLeuAlaLeuProMetSerProSerP 267
   ||||| : : : : : |||||
2196 TACGACCGCATTC.....TTGCC...AGCCGAGGG 2224
   267 roleuSerProSerProIleProSerProAsnAlaLysLeuAsnSer 283
   ||||| ||| ||||| : : : : :
2225 GCGTGTCCAAGTTCCGTGGCATCACCG...AGCAGGCGGAGATCTCT 2271
   284 AlaLeuLeuThrValGluProSerPro..... 292
   ||| ||||| : : : : : |||
2272 CCGCTAGTAGTACGGCAACCCGAAACCGCGCGGCAAACTCGAAGCTCTG 2321
   293 .....GlnAspLysAsnLysGlyPhePheValAspGluSerGluProL 307
   ||| : : : : : ||| : : : : :
2322 GTCTCAACAACGCTTGAGCATCGCTTCGAGCTCCGCACTCACCGA 2371
   307 euleuArgCysAspSerThrSer.....SerGlySerSerAlaLeu 320
   : : : : : : : : : : : : : : : :
2372 TCACCAACGCGCGGTACACATCATGAGTAGATCTTGAAGTCAAGCTTGA 2421
   321 SerArgAsnGlySerPheIleThrLysGluLysLysAspThrValLeu.. 336
   : : : : : : : : : : : : : : : :
2422 AGAAGTACGCGCTCTGTGCACAGAGATCAAGCGTTCAAGTACTTCTC 2471
   337 .....ArgGlnValArgLeuAspProGlyAspLeuGlnProI 349
   ||| ||||| : : : : : |||
2472 ACCCAGTCCAGCGCGGTGACGAGATGATGATCCCTCAACTCGGATCA.. 2520
   349 IePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIle 365
   2521 .....GGCTTCTCCACGCAAGCAGACCTCTCTCGTGGAAGATTTC 2562
   366 GluGlu 367
   ||||| : : : : :
2563 GAGCAA 2568

```

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-324-867-1

```

seq_documentation_block:
  Sequence 1, Application US/09324867A
  Patent No. 6251632
  GENERAL INFORMATION:
  APPLICANT: Lillistrap, David
  APPLICANT: Cameron, Cherie
  APPLICANT: No. 6251632Iley, Colleen
  APPLICANT: Horrocks, L. Suzanne Hoyle
  APPLICANT: Hough, Christine
  TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
  FILE REFERENCE: 1669 0010002/JAG/RJD
  CURRENT APPLICATION NUMBER: US/09/324,867A
  CURRENT FILING DATE: 1999-06-03
  EARLIER APPLICATION NUMBER: 09/035,141
  EARLIER FILING DATE: 1998-03-059
  EARLIER APPLICATION NUMBER: 60/039,953
  NUMBER OF SEQ ID NOS: 63
  SOFTWARE: PatentIn Ver. 2.0
  SEQ ID NO 1
  LENGTH: 7032
  TYPE: DNA
  ORGANISM: Canis familiaris
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 1..7029
  US-09-324-867-1

```

alignment_scores: Quality: 99.00 Length: 458

Ratio: 0.503 Gaps: 19
Percent Similarity: 43.013 Percent Identity: 19.432

Alignment block:

US-09-733-956-2 x US-09-324-867-1

Align seg 1/1 to: US-09-324-867-1 from: 1 to: 7032

```

10 AlaservAlaArgProLysValLeuSerSerLeuGluGlyThrValPr 26
   |||||.....|
2536 GCAAGTCTCAGACGAGCTCGTCACAGATGAGACAGAGAAATTTTCTCC 2585
   |||||.....|
26 AspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeup 43
   |||||.....|
43 roAsnLeuGluVal...ValAsnHisGluGlyProHisHisArgHis 58
   |||||.....|
2591 CAGAACTGCAGCTTAAGATTAAATGAGAAATTTGGGCAAAATACACAGCTA 2640
   |||||.....|
59 IleuLys.....|
2641 GAGTTGAAAGAACTTGATTAAATTTCTAGTTATCAGACAGCTCTAAT 2690
   |||||.....|
63 uProSerMetGluAlaThrGlyGlyLysSerSerThrProLeuLysG 80
   |||||.....|
2691 GACTTCACCAACAATTCATCAGATAGTTGGCAGCAGCTACTGTAAGA 2740
   |||||.....|
80 LyrProLysArgGlyHisProArgGlnAsnLeuHisLysHisPheAsp 96
   |||||.....|
2741 CAGGTTCCCTTAGGACCCCAAAATATGTCAGTTCCAC.....TTT 2778
   |||||.....|
97 AsnGluHisLeuProThrMetLeuVal.....|
2779 AACAGCTCATTTAGTACCATGTTGATTGCAATATTCATCCACCTTAT 2828
   |||||.....|
106 .....|
2829 TCAGTCTGCTGACCTTTGGAATGAGTGAAGAAATATGATTCACAGT 2878
   |||||.....|
107 heLeuLeuLeuValLeuValValIleValValCysSerIleArg..... 121
   |||||.....|
2879 TGTTAAGAGACACCTTAATGATATATCAAGAAATCTCAGTGAAGAAAT 2928
   |||||.....|
122 .....LysSerSerArgThrLeuLysLysGlyProArgGlnAs 134
   |||||.....|
2929 GTATTATCAATGAGAGATTAATGTTATTATAAGAAAGAAATTCGTGG 2978
   |||||.....|
134 proSerAlaIleValGluLys.....|
2979 ACCTGCTTCATTAAATCAAGATTAATGCTTTATTCAAAGTTAATATCTCT 3028
   |||||.....|
143 LysLeuLysLysSerMetThrPro.....ThrGlnAsnArgGluLys 156
   |||||.....|
3029 CGGTAAAGACAAACAGGACGACCTTAACCTTAACAACTAATGAAAGACT 3078
   |||||.....|
157 TrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuValAl 173
   |||||.....|
3079 .....CGTGTGCTATCCCAACATTATTAT 3104
   |||||.....|
173 aAlaGluValGlySerGlnTrpLysAspIleTyrGlnPheLeuCysAsn 190
   |||||.....|
3105 TGAGAACAGTACCTCAGCTGCGCAAGATATTT...ATGTTAGAAAGAAATA 3151
   |||||.....|
190 laSerGluArgGluValAlaAlaPheSerAsnGlyTyrThr...AlaAsp 205
   |||||.....|
3152 CTGAGTTTAAAGAACTAATCTTTGATTCATAAAGAAACGTTTATGCGC 3201
   |||||.....|
206 HisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgGlyProAl 222
   |||||.....|
3202 AGAAATACTACAGCTCTGGGGCTTAATCATGTGTCA.....AAATA 3242
   |||||.....|
222 uAlaSerLeuAlaGluLeuIleSerAlaLeuArgGlnHisArgArgAsn 239

```

```

3243 AACCTATTATATCAAAAAATGCGAAATGCGCCACCAAAAAA..... 3285
   |||||.....|
239 spValValGluLysIleArgGlyLeuMetGluAspThrThrGlnLeuGlu 255
   |||||.....|
3286 .....GAAGACCTGTCCTGCTGCTGCT 3306
   |||||.....|
256 ThrAspLysLeuAlaLeuProMetSerProSerProSerProSerPr 272
   |||||.....|
3307 GCAGAAATCCGAGATCTATCATCTCTCAAGATACCGTTCTTGCCAGATTG 3356
   |||||.....|
272 oLleProSerProAsnAlaLysLeuGluAsnSerAlaLeuLeuThrValG 289
   |||||.....|
3357 GATTAAGAACCCATGGC.....AAGAACTCCCTTAACCTCTGAGCANA 3397
   |||||.....|
289 LuProSerProGlnAspLysAsnLysGlyPhePheValAspLysSerGlu 305
   |||||.....|
3398 GCCCAGTCCTCAAAA..... 3411
   |||||.....|
306 ProLeuLeuArgCysAspSerThrSerSerGlySerSerAlaLeuSerAr 322
   |||||.....|
3412 .....CAATTAAATCTTTAGGATCAGAAATAATCTGTGTA 3446
   |||||.....|
322 gAsnGlySerPheIleThrLysGluLys..... 331
   |||||.....|
3447 ACATCAGAACTTTTGTCTAGAGAGAGAGCTGCTAGTACGAGAGATGAAT 3496
   |||||.....|
332 .....LysAspThrValLeuAlaArgGluValAlaArgLeuAspProCysAspLeu 346
   |||||.....|
3497 TTACGAAAGACACA.....GAACTC 3516
   |||||.....|
347 GlnProIlePhe.....AspAspMetLeuHisPheLeuAsn..... 358
   |||||.....|
3517 CAAGAGATTTTTCACAAACACAGAGACATATTTTGTGTAACCTTGCTGA 3566
   |||||.....|
359 .....ProGluGluL 362
   |||||.....|
3567 TGTCCAGAAATATGATACATCAATCAAGAAAAAATCTCCGGAAGACA 3616
   |||||.....|
362 euArgValIleGluGluIleProGlnAlaGluAspLysLeuAspArgLeu 378
   |||||.....|
3617 TGAAGAAAGAAAGAAATTAACCAAGAGAAATGCTGCTTGGCTCAGGCA 3666
   |||||.....|
379 PheGluIleIleGlyValLysSer 386
   |||||.....|
3667 CATCTATGATTTGGCCTAAGAAC 3690
   |||||.....|
seq_name: /cgn2_6/plodata/2/lna/5b_COMB.seq:US-08-989-925-2
seq_documentation_block:
: Sequence 2, Application US/08989925
: Patent No. 5989820
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Lal, Preeti
: APPLICANT: Shah, Purni
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN ADIPOPHILIN-LIKE PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/989,925

```


OM of: US-09-733-956-2 to: N_Geneseq_1101.* out_format : pfs
Date: Jan 4, 2002 9:11 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framed+g2n.model -DEV=x1p
-O=/cgml_1/USPTO.spool/US09733956/crunat_04012002_084026_15566/app-query.fasta.1.467
-DB=N_Geneseq_1101 -QPM=fastap -SUFFIX=ing -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPL=0.000 -LOOPEXT=0.000
-GAPOP=6.000 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pt
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFORMAT=pfs
-NORTEXT -MINLEN=0 -MAXLEN=200000000
-USER=US09733956_0CGML_1_305 -MCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-733-956-2
Query length: 405
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 163.610000*

score_list:

Sequence	Strid	Orig	ZScore	EScore	len	Documentation
/SID8/gcgdata/geneseq/geneseq/NA1998.DAT:AAV57441					2091.00	3492.71
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV28728					2091.00	3492.29
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV57184					2091.00	3490.53
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV292406					2091.00	3489.69
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV47395					2091.00	3488.39
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV47395					2091.00	3488.25
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25259					2091.00	3487.71
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV99927					2091.00	3487.61
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV99658					2091.00	3487.61
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25322					2091.00	3487.42
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25322					2091.00	3487.42
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV14374					2087.00	3483.91
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25320					1914.00	3190.03
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25264					1617.00	2611.97
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25322					1527.00	2540.96
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV28727					1511.00	2515.09
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV28727					1352.00	2252.69
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25261					1322.00	2195.57
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25321					1269.00	2106.64
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV25260					1269.00	2106.49
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV292407					992.50	1647.36
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV112543					804.00	1342.63
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV133893					804.00	1342.63
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV102448					804.00	1342.63
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV99929					579.00	966.10
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV99660					579.00	966.10
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV99928					564.00	941.23
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV99659					564.00	941.23
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV87597					450.00	751.96
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV121181					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV122711					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV146449					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV148007					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV106386					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV108374					439.00	734.53
/SID8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV99921					398.50	664.15
/SID8/gcgdata/geneseq/geneseq/NA2000.DAT:AAV99662					398.50	664.15
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV122923					397.00	665.22
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV148220					397.00	665.22
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV108587					397.00	665.22

/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV11976	373.00	616.95	2.7e-26	465
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV13308	373.00	616.95	2.7e-26	465
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV10199	373.00	616.95	2.7e-26	465
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV13702	349.00	579.24	3.5e-24	364
/SID8/gcgdata/geneseq/geneseq/NA2001.DAT:AAV135059	349.00	579.24	3.5e-24	364
seq_name: /SID8/gcgdata/geneseq/geneseq/NA1998.DAT:AAV57441				
seq_documentation_block:				
ID AAV57441 standard; cDNA; 2186 BP.				
AC AAV57441;				
XX 21-DEC-1998 (first entry)				
DT Human tumour necrosis factor related receptor TR7 cDNA.				
XX				
XX Tumour necrosis factor related receptor; TR7; human; inflammation;				
KW arthritis; septicemia; autoimmune disease; psoriasis;				
KW inflammatory bowel disease; infection; graft-versus-host disease;				
KW transplant rejection; stroke; acute respiratory disease syndrome;				
KW ischaemia; restenosis; brain injury; AIDS; bone disease; cancer;				
KW atherosclerosis; Alzheimer's disease; diagnosis; therapy; vaccine;				
SS.				
OS Homo sapiens.				
XX				
Key Location/Qualifiers				
FT CDS 7..1974				
FT FT /*tag a				
FT FT /transl_except= (pos:76..78, aa:Arg)				
XX				
PN EP869179-A1.				
XX				
PD 07-OCT-1998.				
XX				
PF 01-APR-1998; 98EP-0302528.				
XX				
PR 28-OCT-1997; 97US-0959382.				
PR 02-APR-1997; 97US-0041796.				
XX				
PI (SMK) SMITHKLINE BEECHAM CORP.				
XX				
DR Deen KC, Hurler MR, Tan KB, Young P;				
XX				
WP1: 1998-508493/44.				
XX				
P-PSDB; AAV75792.				
XX				
PS New tumour necrosis factor receptor TR7 polypeptides and				
XX				
Alzheimer's disease, AIDS and cancer				
XX				
PS Claim 4: Page 18-19; 25pp; English.				
XX				
CC This cDNA clone codes for a novel member of the human tumour				
CC necrosis factor (TNF) receptor superfamily, termed TR7 (see AAV75792).				
CC 3 ESTs with sequence similarity to the human TNF receptor were				
CC discovered in a commercial EST database. Analysis of 2 of these				
CC regions, overlapping, with 100% identity. Together, the 2 clone				
CC sequences encompassed the complete predicted coding sequence of TR7				
CC of 1968 bp. TR7 polynucleotides are useful for diagnosing				
CC diseases or susceptibility to diseases by identifying mutations in				
CC the TR7 gene using probes containing the TR7 nucleotide sequence.				
CC TR7 polypeptide can be used to screen for agonists and antagonists				
CC which bind the receptor. Gene therapy may be used to effect				
CC endogenous TR7 polypeptide production. TR7 polypeptides and				
CC polynucleotides can be administered directly as vaccines for				
CC prevention of diseases. Diseases diagnosed, treated or prevented				
CC by the above methods include: chronic and acute inflammation,				
CC arthritis, septicemia, autoimmune diseases (e.g. inflammatory bowel				
CC disease, psoriasis), transplant rejection, graft vs. host disease,				
CC infection, stroke, ischaemia, acute respiratory disease syndrome.				
CC				

CC restenosis, brain injury, AIDS, bone diseases, cancer (e.g.
 CC lymphoproliferative disorders), atherosclerosis, and Alzheimer's
 CC disease. The TIR polynucleotide is also useful for mapping the
 CC gene to a chromosome, allowing gene inheritance to be studied
 CC through linkage analysis.

XX Sequence 2186 BP; 514 A; 613 C; 565 G; 494 T; 0 other:

alignment_scores:

Quality: 2091.00 Length: 405
 Ratio: 5.163 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-733-956-2 x AAV57441

Align seg 1/1 to: AAV57441 from: 1 to: 2186

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
757 ATGACATCAACAGAAATCAACTCTCTGCTCTGTTAGACCAAGGTAAGTACT 806
17 userSerIleGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
807 GAGTACATCCAGAGAGGACGACGCTCCAGACACACAGCTCAGCAAGG 856
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
|||||
857 GGAGAGAGAGCGTGAACAGACCCCTCCAAACCTTCAGGTAAGTCAACAC 906
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG 67
|||||
907 CAGCAAGGCGCCCAACACACACATCTGAAAGCTCTGCGCTGATGGA 956
67 uAlaThrGlyGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
957 GGGCACTGGGGGGGAGAGACTCCAGCACGCCCATCAAGGGCCCCAAGAGG 1006
84 LysIleProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1007 GACATCTAGACAGAACCTACACACAGCATTTTGACATCAATGAGCATTTG 1056
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValVal 117
|||||
1057 CCTTGATATGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
|||||
1107 GTCCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAGGGGGCCCCGCGCAG 1156
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1157 ATCCAGTGCATTTGTGAAAAGCGAGGCTGAAGAAATCCATGACTGCCA 1206
151 ThrGlnAsnArgGluLysTrpIleTyrrYrcYasngLHisLysIleAs 167
|||||
1207 ACCCAACACCGGGAGAAATGATCTACTCTGCAATGGCCATCTGTATCCA 1256
167 PileLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIle 184
|||||
1257 TATCCCTGAAGCTGTGACAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1306
184 yTcGlnPheLeuCyAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
|||||
1307 ATCAGTTTCTTTCAMTCCAGTGAAGAGAGGTCTGCTGTTCTCCCAAT 1356
201 GlyTyrrThAlaAspHisGluArgAlaTyrrAlaAlaLeuGlnHisTrp 217
|||||
1357 GGGTACACAGCCGACCAAGCGAGCGGCTTACGCACTCTGACGACATGAG 1406
217 rIleArgGlyProGluAlaSerIleuAlaGlnIleuLieserAlaLeuArg 234
|||||

```

```

1407 CATCGGGGCCCCGAGGCCAGCTCGCCAGCTAATTAGCGCCCTCGGCC 1456
234 LHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
|||||
1457 AGCACCAGGAAAGCATGTTGTGGAGAAAGATTCGTGGCTGATGGAGAGC 1506
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuPrometSerProSerP 267
|||||
1507 ACCACCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCGCCAGCC 1556
267 oleuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSer 284
|||||
1557 GCTTACCCGAGCCCATCCAGCCCGCCCAACGGAATTTGAAGATTCGG 1606
284 lAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1607 CTCTCTGACGGTGGAGCCTTCGCCACAGACAGACAGAAAGAGGCTTCTTC 1656
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
|||||
1657 GTGATGAGATCGGAGCCCTCTCTCCGCTGTGACTACATCCAGCGGCTC 1706
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAsp 334
|||||
1707 CTCCGCGCTGAGCAGGACGCTTCCTTTATACCAAGAAAGAAAGAGACA 1756
334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
1757 CAGTGTGGGCGAGTACCGCTGAGCCCTGTGACTTGGACCTATCTTT 1806
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluL 367
|||||
1807 GATGACATCTCCACTTCTTAATCTGAGGAGCTCGGGGTATGGAAGA 1856
367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
|||||
1857 GATTCGCCAGCGTGAAGACAACTACACCGGCTATTGCAATATTATGGAG 1906
384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrrSerHis 400
|||||
1907 TCMAAGAGCCAGAGGACGACAGACCCCTCGGACTGTGTTATACCAT 1956
401 LeuProAspLeuLeu 405
|||||
1957 CTTCCTGACCTCTGCTG 1971

seq_name: /SID58/gcdata/geneseq/geneseq/NA2000.DAT:AAA28728
seq_documentation_block:
ID AAA28728 standard; DNA: 2271 BP.
XX
XX AAA28728;
XX
XX 29-AUG-2000 (first entry)
XX
XX DE Human osteoprotegerin-like 2 (OPGx2) DNA.
XX
XX OPGx1: osteoprotegerin-like; DR6 TNF-related death receptor; agonist;
XX inhibitor; bone resorption; vascular calcification; apoptosis;
XX osteopethic; chromosome 6p11.1; D6S452; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX CDS 102..2177
XX FT /*tag= a
XX FT /product= OPGx2
XX
XX WO200024771-A2.
XX
XX 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US24913.
XX
XX PF

```

XX 23-OCT-1998; 98US-0105481.
PR 01-OCT-1999; 99US-0156993.
PR 21-OCT-1999; 99US-0422680.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Yang M, Lichenstein H, McDonald WF;
XX WPI: 2000-350692/30.
DR P-PSDB; AAF92846.
XX
PT Nucleic acids encoding osteoprotegerin-like (OPGx) polypeptides, useful
PT for treating disorders associated with bone metabolism, such as
PT osteoporosis and osteopetrosis
XX
PS Claim 1; Fig 2; 81pp; English.
XX
CC Osteoprotegerin-like polypeptides (OPGx) are expressed in bone, lymph
CC node, germinal B cells and kidney. There appear to be at least two splice
CC variants of this gene, a transmembrane form and an extracellular domain
CC form. Human OPGx1, 2 and 4 contain a 36 amino acid extension at their
CC N-terminal relative to the DR6 tumour necrosis factor (TNF)-related death
CC receptor. The OPGx gene has been mapped to chromosome 6p11.1, essentially
CC on top of D6S452. The OPGx polypeptides, agonists and antibodies are
CC useful in methods to inhibit osteoclast-mediated bone resorption or
CC vascular calcification and to modulate cell death (apoptosis). This is
CC useful for treating disorders associated with bone metabolism, such as
CC osteoporosis, osteopetrosis, or a condition characterized by loss of
CC bone, breakdown of tissue, or excessive readorption of bone tissue.
XX
XX Sequence 2271 BP: 518 A; 668 C; 611 G; 467 T; 7 other:

alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AAA28728 ..

Align seg 1/1 to: AAA28728 from: 1 to: 2271

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProIleValle 17
|||||
960 ATGAACTCAACAGATCCAACTCTCTGCTCTGTAGACCAAGGACT 1009
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
1010 GAGTACATCCAGAGGAGGAGTCCCTGACACACACAGCTCAGCAAGG 1059
34 IyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
|||||
1060 GGAAGGAAAGCGTGAACAGACCCCTCCAAACCTTCAGGTAAGTCAACAC 1109
51 GlnGlnGlyProHisIleArgHisIleLeuLysLeuLeuProSerMetG 67
|||||
1110 CACGACAGGCCCCACCAACACATCTGAACTGCTGCTGCTCAGCA 1159
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 84
|||||
1160 GGGCACTGGGGGGGAGAGTCCGACGCGCCATCAAGGGGCCCAAGAGGG 1209
84 IyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1210 GACATCTAGACAGAACTACACACACATTGTCATCAATGAGCATTTTG 1259
101 ProTrMetIleValLeuPheLeuLeuValLeuValIleValIleVal 117
|||||
1260 CCTGGATGATGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309
117 IcySerIleArgLysSerSerArgThrLeuLysGlyProArgGlnA 134

|||||
1310 GTGCAGATATCCGGAAAAAGCTCGAGGACTGTGAAAAAGGGGCCCGGACAG 1359
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1360 ATCCAGATGCCATTGTGGAAAAAGGAGGCTGGAAGAAATCCATGACTCCA 1409
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
|||||
1410 ACCCAGACCGGGAGAAATGATCTACTACTGCAATGGCCATGCTATCCA 1459
167 PileLeuLysLeuValAlaIleValGlnValGlySerGlnTrpLysAspIleT 184
|||||
1460 TATCTGGAAGCTTGTAGACACCAAGTGGGACCAAGTGGAAAGATATCT 1509
184 yTrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
|||||
1510 ATCAGTTTCTTTTCCAAATGCCAGTACAGAGGAGGCTGCTGCTTCTCCAA 1559
201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
|||||
1560 GGGTACACAGCCGACCAAGCGGGGCTTACGAGCTCTGACGACTGGAC 1609
217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||
1610 CATCCGGGGGCCCGAGGCCAGCTCGGCCAGTAATTAGCGCCCTGGCC 1659
234 IyHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
|||||
1660 AGCACCAGGAGAAACGATGTTGTGAGAAAGATTCGTGGCTGATGGAAGAC 1709
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1710 ACCACCCAGCTGGAACCTGACAACTAGCTCTCCGATGAGCCCGACCC 1759
267 OLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSera 284
|||||
1760 GCTTAGCCCGGAGCCCATCCCGACGCCCAAGCCAACTGGAATAATTCCG 1809
284 IAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1810 CTCTCTGACGGTGGAGCCCTTCCACACAGGACAAAGAAAGAGGCTTCTTC 1859
301 ValAspGluSerArgLupProLeuLeuArgCysAspSerThrSerSerGly 317
|||||
1860 GTGATGATGATGCGAGCCCTTCTTCCGCTGTGACTCTACATCCAGCGCTC 1909
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAsp 334
|||||
1910 CTCGCGGCTGAGCAGGAAAGGTTCTTTATTTACCAAGAAAGAAAGAGCA 1959
334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
1960 CAGTGTGGCGGAGGTAGCGCTTGAGCCCTGTGACTGTGACGCTTACTCTT 2009
351 AspAspMetLeuHisPheLeuAsnProGlnGluLeuArgValIleGluG 367
|||||
2010 GATGACATGCTCCACTTTCTAAATCTGAGGAGCTGGGGGTGATTGAGA 2059
367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
|||||
2060 GATTCCCGCAGCGTGAAGACAACTAGACCGGCTATTGCAAAATTATTCGAG 2109
384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
|||||
2110 TCAAGAGCCGAGGAAGCCAGACACCTCTCTGAGACTCTGTTTATAGCCAT 2159
401 LeuProAspLeuLeu 405
|||||
2160 CTTCCTGACTGCTG 2174
seq_name: /SID8/gcdata/geneseq/geneseqn/NA2000.DAT:AA257184

seq_documentation_block:
ID AA257184 standard; CDNA: 2666 BP.
AC AA257184;
DT 27-MAR-2000 (first entry)
XX
DE Human DETH protein encoding CDNA SEQ ID NO:1.
XX
KW Human; DETH; apoptosis; TNFR; neurodegenerative diseases; cancer;
KM death domain expressing tumor necrosis factor receptor family homologue;
KW autoimmune disease; diagnosis; ss.
XX
OS Homo sapiens.
XX
PN M09966039-A1.
XX
PD 23-DEC-1999.
XX
PF 08-JUN-1999; 99MO-GB01793.
XX
PR 12-JUN-1998; 98GB-0012607.
XX
PA (ZENE) ZENECA LTD.
XX
PI Lu JJ, Gomes BC, Fieles WE;
XX
DR WPI: 2000-097744/08.
DR P-PSDB: AAY67947.
XX
PT New protein having 80% identity to death domain expressing tumor
PT necrosis factor receptor family homologue -
XX
PS Claim 7; Page 33-34; 40pp; English.
XX
CC The present sequence encodes a human death domain expressing tumor
CC necrosis factor receptor family homologue, designated DETH. DETH has a
CC role in apoptosis and appears to be member of the TNFR1/FAS/NGFR family
CC of receptors. The DETH protein can be used for inducing apoptosis by
CC expressing DETH in a cell. The protein is useful for identifying
CC inhibitors and agonists of apoptosis that treat neurodegenerative
CC diseases (inhibition); cancer and autoimmune diseases (induction).
CC DETH-specific antibodies are useful for diagnosis of conditions and
CC diseases associated with DETH expression. The protein avoids the use
CC of harmful chemotherapy, where the loss of the p53 tumour repressor
CC gene can lead to drug resistant tumour cells following treatment.
XX
SQ Sequence 2666 BP; 691 A; 609 C; 591 G; 775 T; 0 other;

alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AA257184 ..

Align seg 1/1 to: AA257184 from: 1 to: 2666

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
190 ATGAACCTCAACAGAAATCCAACTCTTCTGCTGTAGACCAAAAGTACT 239
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
240 GAGTAGCATCCAGGAAGGACAGTCCCTGTACAAACAACTAGCAGCAGG 289
34 IlyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
|||||
290 GGAAAGAGACGCTGTAACAAGACCTCCCAAACTTCAGGTAAGTCAACAC 339
51 GlnGlnGlyProHisIleArgHisIleLeuLysLeuLeuProSerMetGcl 67

|||||
340 CAGCAAGGCCCCACCACAGACATCCGAAAGCTGCTCCATCGA 389
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 84
|||||
390 GGCACCTGGGGGCGAGAACTCCACACGCCCTCAAGAGGCCCAAGAGG 439
84 IyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
|||||
440 GACATTCCTAGACAGAACCTACACAGCATTTTGACATCAATGAGCATTTG 489
101 ProTPMetIleValLeuPheLeuLeuValIleValIleValIleVal 117
|||||
490 CCTGGATGATTTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
117 IcySerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
|||||
540 GTGCAGTATCCGGAAAGCTCGAGGACTGTGAAAAGGGCCCGGCGAG 589
134 sProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
590 ATCCCATGTCATTTGTGMAAGCGAGGCTGAAAGAAATCCATGACTCCA 639
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
|||||
640 ACCCAGAACCGGAGAAATGATCTACTACTCAATGGCATGATATCGA 689
167 ILeuLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
|||||
690 TATCTGTAAGCTTTGTAAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 739
184 yTrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
|||||
740 ATCAGTTTCTTGGCAATGCCAGTAGAGAGGAGTTGCTCTTCTTCATAT 789
201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
|||||
790 GGTATCACAGCCGACACAGCGGCGCTACGAGCTCTGCACACTGGAC 839
217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||
840 CATCGGGGGCCCGAGGCGAGCTCGCCAGCTAATTAGCGCCCGCGCC 889
234 IHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
|||||
890 AECACCGGAGAAACGATGTTGGAGAGATTCGTGGCTGATGGAAAC 939
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
940 ACCACCCAGCTGGAACTGACAACTAGCTCTCCGATGAGCCCAAGCC 989
267 OleuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSer 284
|||||
990 GCTTAGCCCGAGCCCATTCGCCAGCCCAAGCGAAACTGGAATTCG 1039
284 ILeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1040 CTCTCTGACGGTGGAGCTTCCCAAGAGCAAGAAACAGGCTTCTTCC 1089
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGly 317
|||||
1090 GTGGATGAGTGGAGCCCTTCTCCGCTGTGACTCTACATCCAGGCGCTC 1139
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysAspT 334
|||||
1140 CTCGCCGCTGAGACAGAAAGCTTCTTATTATTCAAAGAAAGAAAGACA 1189
334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
1190 CAGTGTGGGCGAGTAGCGCTGAGACCCCTTGACTGACTGAGCTATCTTT 1239
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGlnG 367
|||||

```
1240 GATGACATGCTCCACTTCTTAATCTGAGAGCTCGGGTGATTGAAGA 1289
367 uilleproglinalagluaspysleuaspargleuphegluilellely 384
1290 GATTCCCGACGCTGAGACAAACTAGACCGCTATTTGAAATTAATTGAG 1339
384 allysserlinglualeserglthrleuenuaspservaltyrserhis 400
1340 TCAAGAGCCAGGAAGCCAGACACCTCTGAGACTCTGTTTATAGCAT 1389
401 LeuproAspleuLeu 405
1390 CTTCCTGACCTGCTG 1404

seq_name: /STDs/gcgdata/geneseq/geneseqn/NA2000.DAT.AA292406
seq_documentation_block:
ID AA292406 standard; cDNA: 2877 BP.
XX
AC AA292406;
XX
DT 05-JUN-2000 (first entry)
XX
DE cDNA encoding human TNF receptor-like protein HSLJD37R. SEQ ID NO:9.
XX
KM TNF receptor family: tumour necrosis factor; HDTEA84; HSLJD37R;
KM Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KM immunosuppressant; antirheitic; antirheumatoid; antiinflammatory;
KM dermatological; antithyroid; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 410..2377
FT /tag= a
FT /product= "Human HSLJD37R"
XX
PN WO200001817-A2.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99WO-US12366.
XX
PR 06-JUL-1998; 98US-0110938.
PR 13-JUL-1998; 98US-0114466.
PR 23-JUL-1998; 98US-0093897.
PR 12-AUG-1998; 98US-0132968.
PR 18-AUG-1998; 98US-0136214.
PR 11-SEP-1998; 98US-0099999.
XX
(SCHER) SCHERING CORP.
XX
PI Bates EM, Lebeque SJF, Murphy EE, Mattson JD, Gorman DM;
PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
PI Bazan JF, Mahony D, Lees EM;
XX
DR MPI: 2000-171015/15.
DR P-PSDB: AAY77460.
XX
PT New isolated mammalian genes, used to develop products for treating
PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions -
XX
PS Claim 25; Page 160-163; 218pp; English.
XX
CC The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC human CC chemokine RCC5; human deubiquitinating proteins Dab1 and Dub
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif
CC (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
```

```
CC against these proteins. The proteins can be used for modulating the
CC physiology or development of a cell. They can be used to mediate uptake
CC of substrates (e.g., prostaglandin-like molecules), to modulate or
CC mediate cellular interactions (e.g., induce or prevent trafficking,
CC proliferation, or differentiation of cells), or are intracellular
CC proteins which are important in various cellular processes such as the
CC deubiquitination of proteins or cell cycle regulation. The products can
CC be used for treating medical conditions such as immune, inflammatory or
CC allergic disorders, or abnormal cellular proliferation, for example,
CC cancers or degenerative conditions. They can be used to modulate immune
CC responses in disease states e.g., autoimmune disorders, including
CC rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC autoimmune thyroiditis, as well as acute and chronic inflammatory
CC responses in which T cell activation, expansion, and/or immunological T
CC cell memory play an important role. Sequences AA292404-292411 represent
CC cDNAs encoding TNF receptor family-related proteins. AA292404 encodes the
CC human protein HDTEA84, AA292405-292407 encode human HSLJD37R proteins,
CC AA292408 encodes murine Rank-like protein RANKL, and AA292409-292411
CC encode human RANKL proteins.
XX
XX
SO Sequence 2877 BP; 636 A; 817 C; 751 G; 673 T; 0 other;
```

```
alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-733-956-2 x AA292406 ..
```

```
Align seg 1/1 to: AA292406 from: 1 to: 2877
```

```
1 MetasnerThrGluSerAsnSerSerAlaSerValArgPolyValle 17
1160 ATGACACTCACAGAACATCCAACTTCTGCTGCTGTAGACCAAGGACT 1209
17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
1210 GAGTAGCACTCCAGAGAGGACAGTCCCTGACACACACAGCTCAGCAGAGG 1259
34 lyLysGluAspValAlaAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
1260 GGAAGGAAGACGTGACAGACACCTCCCAACCTTCAGCTGTCACACAC 1309
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG1 67
1310 CAGCAAGGCCCCACACACAGACATCCTGAAGCTGCTGCGCTCCATGGA 1359
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
1360 GCCCACTGGGGGCGAAGATCCACAGCCCATCAAGGGCCCAAGAGAGG 1409
84 LysHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
1410 GACATCTAGACAGAACCTACCAAGCATTTTGACATCAATGACAGCATTTG 1459
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValIleVal 117
1460 CCCGTGATGATTTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
1510 GTGCAGTATCCGGAAGACCTGAGGACCTGAAAAGAGGGCCCGCGCAGG 1559
134 sProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
1560 ATCCAGATGCATTGTGAAAAGGAGGCTGAGAAGAAATTCATGACTCCA 1609
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
1610 ACCCAGAACCGGGAAGAAATGATCTACTACTGCAATGGCCATGGTATGCA 1659
```

```

167 pilleuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleR 184
1660 TATCCGGAAGCTTTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1709
184 yfGlnPheLeuCyAsnAlaSerGluArgGluValAlaIlePheSerAsn 200
1710 ATCAGTTTCTTTGCAATGCCAGTGAAGAGGAGGTTCCTCTCTTCTCAAT 1759
201 gLyTrpThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
1760 GGGTACACAGCCGACGACGAGGGGCTTACGACGCTCTCAGACACTGGAC 1809
217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
1810 CATCCGGGGCCCGGAGCCAGCTCCGCGCAATTAATTAACGCGCTCGGCC 1859
234 lHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
1860 AGCACCAGGAAGAACGATGTTGTGGAGAGATTCGTGGCTGATGAGAGAC 1909
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
1910 ACCACCCAGCTGGAACGACAACTAGCTCTCCGATGATGAGCCCGCC 1959
267 OleuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
1960 GCTTATCCCGAGCCCATCCATCCGAGCCCAAGCAACTTGAAGATTCGG 2009
284 lLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
2010 CTCTCTGACGGTGGAGCCTTCCCCACAGGACAGAACAGGCGTTCTTC 2059
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
2060 GAGGAGATCGAGAGCCCTTCTCCGCTGACTTACATCCAGCGCGCTC 2109
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
2110 CTCCTGGCTGAGCAGGAGACGGTCTTATATACCAAGAAAGAAAGGACA 2159
334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
2160 CAGTGTGGGCGAGTACGCTCGACCTGTGACTTGACAGCTATCTTT 2209
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluG 367
2210 GATGACATCTCTCACTTCTAAATCCTGAGAGAGCTCGGGTGATTGAGA 2259
367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleGlyLV 384
2260 GATTCCTCCAGGCTGAGGACAACTAGCCGCGCTATTGAAATTTATGGAG 2309
384 allySerGlnGlnAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
2310 TCAAGAGCCAGAGAGCCAGCCAGCCCTCTGACTCTTTATATACCANT 2359
401 LeuProAspLeuLeu 405
2360 CTTCCTGACCTGCTG 2374
seq_name: /SIDS8/gcdata/geneseq/geneseqn/NA2000.DAT:AAA47396
seq_documentation_block:
ID AAA47396 standard; cDNA: 3236 BP.
XX
AC AAA47396;
XX
XX 20-OCT-2000 (first entry)
XX
XX Tumour necrosis factor receptor homologue TRH1 (Clone 2098183).
XX
XX Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW transplant rejection; activation; proliferation; differentiation;

```

```

KW apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW probe; primer; human; ds.
XX
OS Homo sapiens.
XX
EN W0200034294-A2.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99W0-US29400.
XX
PR 11-DEC-1998; 98US-0111826.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Bowen MA, Siemers N;
XX
DR WPI: 2000-423364/36.
XX
PT Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT development
XX
PS Disclosure: Fig 1b; 42pp; English.
XX
CC The tumour necrosis factor receptor homologue TRH1 can be used for
CC treating a mammal e.g. a human, at risk for a disorder characterized
CC by an aberrant or unwanted level or biological activity of TRH1,
CC e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC useful to leach out or block a ligand which is found to bind to the
CC TRH1. TRH1 may be used in various drug screening techniques and to
CC identify fragments and analogs of a protein or peptide (agonist or
CC antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC cellular function, cell activation, proliferation, differentiation,
CC and apoptosis. The interaction between the novel TNF protein of the
CC present invention and intracellular signaling molecules and/or its
CC potential co-receptor may serve as a novel target for
CC immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC development. Gene constructs can also be used as part of a gene
CC therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC directed against TRH1 can be used to reject TRH1 in tissues
CC and cells. They can also be used to make targeted antibody that
CC destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC used as diagnostic probes or as PCR primers. Fragments of the full
CC length gene may be used as hybridization probes for a cDNA library to
CC isolate the full length gene and to isolate other genes which have a
CC high sequence similarity. The probes may be used to identify a cDNA
CC clone corresponding to a full length transcript and a genomic clone
CC or clones that contain the complete gene including regulatory and
CC promoter regions, exons, and introns.
XX
SQ Sequence 3236 BP; 810 A; 808 C; 751 G; 867 T; 0 other;

```

```

alignment_scores:
Quality: 2091.00      Length: 405
Ratio: 5.163          Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```
alignment_block:
US-09-733-956-2 x AAA47396 ..

```

```
Align seg 1/1 to: AAA47396 from: 1 to: 3236

```

```

1 MetAsnSerThrGluSerAsnSerAlaSerValArgProLysValle 17
|||||
765 ATGAACTCAACAGAAATCAACTCTCTGCTGTAGACCAAGGTACT 814
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
815 GAGTACATCCAGGAAGGACAGTCCCTTACACACACAACTAGCAGAGG 864

```



```

34 1yLysGIuAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
    |||
865 GGAAGGAAGACGTACACAGACCTCCCAAAACCTTCAGGTAGTACACAC 914
    |||
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG1 67
    |||
915 CAGCAAGGCCCCACACAGACATCTCGAAGCTGCGCTCCTCATGGA 964
    |||
67 uAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
    |||
965 GGCCTAGTGGGGGAGAAAGTCCACAGACCCCATAAAGGCCCAAGAGG 1014
    |||
84 LysIleProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
    |||
1015 GACATCTCAGACAGACCTACACAGACATTTGACATCAATGACATTTG 1064
    |||
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValVal 117
    |||
1065 CCGTGGATGATGTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1114
    |||
117 LysSerIleLeuArgLysSerSerArgThrLeuLysGlyProArgGln 134
    |||
1115 GTGCGATGATCGGAAAGCTCGAGAGCTGAAAAAGGGGCCCGGCAAG 1164
    |||
134 sPProSerAlaIleValGlyLysAlaGlyLeuLysLysSerMetThrPro 150
    |||
1165 ATCCCACTGGCATTTGTGAAAGGAGGCGCTGAAGAAATCCATGACTCA 1214
    |||
151 ThrGlnAsnArgGlyLysTrpIleTyrTyrCysAsnGlnHisGlyIleAs 167
    |||
1215 ACCCAGAAACCGGAGAAATGATGATCTACTACGCAATGCCATGTATCGA 1264
    |||
167 PileLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
    |||
1265 TATCTGTAGCTGTAGAGCCCAAGTGGAAAGCCAGTGGAAAGTATCT 1314
    |||
184 YrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
    |||
1315 ATCAGTTTCTTTCATGTCAGTGGAGGAGGTGGCTTCTTCCCAT 1364
    |||
201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
    |||
1365 GGGTACACAGCGCACAGAGCGGCGCTACGACCTGCGACACTGAC 1414
    |||
217 rIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
    |||
1415 CATCCGGGGCCCGAGGCGAGCTCGCCACCTCAATTAGCCCGCTGCGCC 1464
    |||
234 LnhIleArgArgAsnAspValValGlyLysIleArgGlyLeuMetGluAsp 250
    |||
1465 AGCACCGGAGAAAGATGTTGTGGAAAGATTCGTGGCGCTGATGGAAGAC 1514
    |||
251 ThrThrGlnLeuGlnThrAspLysLeuAlaLeuProMetSerProSerPr 267
    |||
1515 ACCACCCAGCGTGAACAGACAACTAGCTCTCCGATGACCCCGACGCC 1564
    |||
267 oLeuSerProSerProIleProSerProAsnAlaLysLeuLysAsnSerA 284
    |||
1565 GCTTAGCCCGAGCCCATCCCAAGGCCCAAGCAACTTGAAGATTCG 1614
    |||
284 lAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
    |||
1615 CTCCTCGAGCGGTGAGCCTTCCACAGACGACAAAGAACAAAGGCTTCTTC 1664
    |||
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
    |||
1665 GTGATGATGTCGAGACCCCTTCCTCGCTGACTCTACATCAGCGGCTC 1714
    |||
317 rSerAlaLeuSerArgAsnGlnLysPheIleThrLysGlyLysLysAspT 334
    |||
1715 CTGCGCGCTGACAGAAAGGATCTTATTATCAAGAAAGAAAGAGGACA 1764
    |||
334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350

```

```

1765 CAGGTGTCGGCAGGTAGCGCTGGACCCCTGTGACTTGACACCTATCTTT 1814
    |||
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluG1 367
    |||
1815 GATGACATGCTCCACTTCTTAATCTGAGAGAGCTGCGGGTGAATTGAAGA 1864
    |||
367 uIleProGlnAlaGlnAspLysLeuAspArgLeuPheGluIleIleGlyV 384
    |||
1865 GATTCCCGAGCTGAGACAAACTAGACCGGCTATTCCAAATTTATTGGAG 1914
    |||
384 aLysSerGlnGlnAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
    |||
1915 TCAAGAGCCAGGAAGCCAGCACACCTCTCGACTCTGTTTATAGCCAT 1964
    |||
401 LeuProAspLeuLeu 405
    |||
1965 CTTCCTGACTGCTG 1979
    |||
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA47395
seq_documentation_block:
ID AAA47395 standard; cDNA; 3277 BP.
XX
XX AAA47395;
AC
XX
XX 20-OCT-2000 (first entry)
DT
XX
XX Tumour necrosis factor receptor homologue TRH1 (Clone 273317).
DE
XX
XX Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW Transplant rejection; activation; proliferation; differentiation;
KW apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW probe; primer; human; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 52.2019
FT CDS /*tag= a
FT /product= Tumour necrosis factor receptor homologue
XX
XX WO200034294-A2.
XX
XX 15-JUN-2000.
PD
XX
XX 10-DEC-1999; 99WO-US29400.
PF
XX
XX 11-DEC-1998; 98US-0111826.
PR
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA
XX
XX Bowen MA, Siemers N;
PI
XX
XX MPI; 2000-423364/36.
DR P-PSDB; AAB01349.
XX
XX Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT development
XX
XX Claim 2; Fig 2a-2d; 42pp; English.
PS
XX
XX The tumour necrosis factor receptor homologue TRH1 can be used for
CC treating a mammal e.g. a human, at risk for a disorder characterized
CC by an aberrant or unwanted level or biological activity of TRH1,
CC e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC useful to leach out or block a ligand which is found to bind to the
CC TRH1. TRH1 may be used in various drug screening techniques and to
CC identify fragments and analogs of a protein or peptide (agonist or
CC antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC cellular function, cell activation, proliferation, differentiation,
CC and apoptosis. The interaction between the novel TNFR protein of the

```

CC present invention and intracellular signaling molecules and/or its
 CC potential co-receptor may serve as a novel target for
 CC immunosuppressive, antiinflammatory and/or immunostimulatory drug
 CC development. Gene constructs can also be used as part of a gene
 CC therapy protocol to deliver nucleic acids encoding the TRH1, or an
 CC agonist or antagonist form of a TRH1 protein or peptide. Antibody
 CC directed against TRH1 can be used to reject TRH1 in tissues
 CC and cells. They can also be used to make targeted antibody that
 CC destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
 CC used as diagnostic probes or as PCR primers. Fragments of the full
 CC length gene may be used as hybridization probes for a cDNA library to
 CC isolate the full length gene and to isolate other genes which have a
 CC high sequence similarity. The probes may be used to identify a cDNA
 CC clone corresponding to a full length transcript and a genomic clone
 CC or clones that contain the complete gene including regulatory and
 CC promoter regions, exons, and introns.

xx
 50 Sequence 3277 BP: 812 A; 825 C; 765 G; 875 T; 0 other:

alignment_scores:

Quality: 2091.00 Length: 405
 Ratio: 5.163 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-733-956-2 x AAA47395 ..

Align seg 1/1 to: AAA47395 from: 1 to: 3277

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProIysValle 17
 ||||||||||||||||||||||||||||||||||||||||||||||||
 802 ATGAGCTCAACAGAAATCCAACTCTTCTGCTCTGTTAGACCAAGGACT 851
 17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
 ||||||||||||||||||||||||||||||||||||||||||||||||
 852 GAGTACATCTCAGAGAGGAGGACACTCCCTGACACACAGCTCAGCAAGG 901
 34 LylsGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 902 GCAAGGAGAGCGTGAACAGACCCCTCCAAACCTTCAGGTAGTCAACAC 951
 51 GlnGlnGlyProHisArgHisIleLeuLysLeuLeuProSerMetG 67
 ||||||||||||||||||||||||||||||||||||||||||||||||
 952 CACGAAAGGCCCCACACACACATCTGAAAGCTGCTGCCGTCATGGA 1001
 67 uAlaThrGlnGlyLysSerSerThrProIleLysGlyProLysArg 84
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1002 GGCCACTGGGGGAGAGAGTCCAGCACGCCCATCAAGGGCCCCAAGAGG 1051
 84 LylHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1052 GACATCTAGACAGAACCTACACACATTTTGCATCATGAGCATTTTG 1101
 101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValIle 117
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1102 CCTGGATGATTTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
 117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1152 GTGCACTATCCGAGAAAGCTCGAGCACTCGAAAGAGGGCCCCGGCAGG 1201
 134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1202 ATCCCACTGTCATTTGGAAAGAGGAGGCTGAAAGAAATCATGAGTCCA 1251
 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1252 ACCCAAGACCGGAGAAATGATCTACTACTGCAATGGCCATGGTATCGA 1301
 167 pIleLeuLysLeuValAlaIleValGlySerGlnTrpLysAspIleT 184
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1302 TATCTGGAAGCTTGTAGCAGGCCCAAGTGGGAGGAGCCAGTGGAAAGAT 1351

184 yrlGlnPheLeuCysAsnAlaSerGluArgGluValAlaIlePheSerAsn 200
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1352 ATGAGTTTCTTTTCATTCAGTCAGAGAGAGAGGAGTCTGCTTTCTCCAA 1401
 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaIleAlaLeuGlnHisTrp 217
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1402 GGGTACACAGCCCGACACAGAGCGGCTTACGAGCTCTCCAGACTGGAC 1451
 217 rIleArgGlyProGlnAlaSerLeuValGlnLeuIleSerAlaLeuArg 234
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1452 CATCCGGGGGCCCGAGGCCAGCTCCGCCAGCAATTTACGGCCCTCGGC 1501
 234 LlnHisArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1502 AGCACCGAGAAACGATGTTGTGAGAAAGATTCGTGGCTGATGGAAGAC 1551
 251 ThrThrGlnLeuGluThrAspLysLeuAlaIleLeuProMetSerProSer 267
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1552 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCCAGGCC 1601
 267 oleUserProSerProIleProSerProAsnAlaLysLeuGluAsnSera 284
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1602 GCTTACCCCGAGGCCCATCCCGCCAGCCCAACGGGAACATTGAGAAATTCG 1651
 284 lAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1652 CTCTCCTGACGCGTGAGACCTTCCCCACAGGACAAAGCAAGGCGCTTTC 1701
 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerLys 317
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1702 GTGATGATGAGTGAGAGCCCTTCTCCGCTGTGACTTACATCCAGCGGCTC 1751
 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLysAsp 334
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1752 CTCGGGGCTGAGCAGGAGACGGTCTTTATTACCAAGAAAGAAAGAGACA 1801
 334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1802 CAGTGTGGGGGAGGTACGCTGGACCCCTGTGACTTGGACGCTTACTT 1851
 351 AspAspMetLeuHisPheLeuAsnProGlnGluLeuArgValIleGln 367
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1852 GATGACATGCTCCACTTTCTAAATCTTGAGAGACTCGGGGTATTTAGAA 1901
 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1902 GATTCCCGAGCGTGAGACAACTAGACCGGCTATTGCAATTTATTTGAG 1951
 384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1952 TCAAGAGCCAGAGAGCCAGACAGACCCTCTGAGACTCTGTTTATACCAT 2001
 401 LeuProAspLeuLeu 405
 ||||||||||||||||
 2002 CTTCCTGACCTGCTG 2016

seq_name: /SID8/gcgdata/geneseq/geneseqn/NA1999.DAT: AAX25259

seq_documentation_block:

ID AAX25259 standard; cDNA, 3440 BP.

AAX25259;

19-JUL-1999 (first entry)

Human full-length tumour necrosis factor receptor ZTNFR-6 cDNA.

ZTNFR-6: tumour necrosis factor receptor-6; human;

cell maturation; bone cell regulation; ss.

Homo sapiens.

XX

```

FH Key Location/Qualifiers
FT CDS 284..2179
FT /tag= a
FT /product= "membrane-bound ZTNFR-6"
FT misc-feature 284..2225
FT /tag= b
FT /note= "this region of the sequence is
FT specifically claimed in Claim 10(a)"
FT sig_peptide 284..334
FT /tag= c
FT mat_peptide 335..2176
FT /tag= d

XX MO9911790-A1.
XX 11-MAR-1999.
XX
XX 03-SEP-1998; 98WO-US18364.
XX
XX 04-SEP-1997; 97US-0923725.
XX 04-SEP-1997; 97US-0057608.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Farrah TM, Gross JA, Mathews SM;
XX
XX WPI: 1999-205190/17.
XX P-PSDB: AAY05678.
XX
XX New secreted or membrane bound tumor necrosis factor receptor
XX ZTNFR-6 - useful for detecting a genetic abnormality in a patient
XX
XX Claim 10a: Page 108-113; 145pp; English.
XX
XX This is the DNA sequence of an isolated polynucleotide that codes
XX for a novel human full-length membrane-bound tumor necrosis factor
XX receptor, designated ZTNFR-6 (see AAY05678). The polynucleotide was
XX initially identified by querying an EST database for sequences
XX homologous to conserved motifs within the TNFR family. A second
XX clone (see AA25260) encodes truncated, soluble ZTNFR-6 (see AAY05679).
XX The full-length clone includes a 40 bp exon (see AA25262) that puts
XX the sequence in frame for the transmembrane and cytoplasmic domains
XX of membrane-bound ZTNFR-6. The human ZTNFR-6 gene was localised to
XX 6p21.1. The isolated polynucleotides can be utilised in the
XX recombinant production of ZTNFR-6 polypeptides, and also in a
XX claimed method for detecting a genetic abnormality in a patient.
XX ZTNFR-6 polypeptides are useful in methods that promote cellular
XX maturation and bone cell regulation.
XX
XX Sequence 3440 BP; 827 A; 876 C; 833 G; 904 T; 0 other;
XX
XX alignment_scores:
XX Quality: 2091.00 Length: 405
XX Ratio: 5.163 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment block:
XX US-09-733-956-2 x AA25259
XX
XX Align seq 1/1 to: AA25259 from: 1 to: 3440
XX
XX 1 MetasnrThrglnuSerAsnSerSerAlaSerValArgProlyValle 17
XX |||||||
XX 962 ATGAACTCAACAGATCCAACTCTTCTGCTCTGTAGACCAAGGTA 1011
XX |||||||
XX 17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaarg 34
XX |||||||
XX 1012 GAGTAGCATCCAGAGGAGGACATCCCTGACACACAAAGCTCAGCAAGG 1061
XX |||||||
XX 34 lylusGluAspValAlaAsnlyThrleuProAsnleuGlnValAlaAsnHis 50
XX |||||||
XX 1062 GGAAGGAAGACGTGAACAGACCTCCCAACCTTCAGTAGTCAACAC 1111

```

```

51 GlnGlnGlyProHisArgHisIleleuLysleuLeuProSerMetG1 67
XX |||||||
1112 CAGCAAGGCCCCACACAGACATCTGTAAGCTGCTGCGCTCATGTA 1161
XX |||||||
67 uAlaThrGlyGlyLysSerSerThrProIleGlyProLysArg 84
XX |||||||
1162 GGCCACTGGGGCGAAGTCCAGACAGCCCATCAAGGGCCCAAGAGG 1211
XX |||||||
84 lylHisProArgGlnAsnleuHislyshIsphAspIleAsnGluHisleu 100
XX |||||||
1212 GACATCCAGACAGAACCTACACAAGCATTTTGACATCAATGACATTG 1261
XX |||||||
101 ProTrpMetIleValleuPheleuLeuValleuValIleValIleVal 117
XX |||||||
1262 CCCTGGATGATGTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311
XX |||||||
117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
XX |||||||
1312 GTGCAGATATCCGGAAGAAAGCTCGAGAGACTGTGAAAAGGGCCCGC 1361
XX |||||||
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
XX |||||||
1362 ATCCAGATGCTATGCGAAAAGCGAGGCTGAAGAAATCCATGACTCCA 1411
XX |||||||
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
XX |||||||
1412 ACCCAGAACCCGGGAAMATGATCTACTACTGCAATGGCCATGATATGA 1461
XX |||||||
167 pIleuLysleuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
XX |||||||
1462 TATCCTGAAAGCTTGATAGCAGCCCAAGTGGAAGCCAGGAAAGATATCT 1511
XX |||||||
184 YrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
XX |||||||
1512 ATCAATTCTTTGCAATCCAGTGAAGAGGAGTGTGCTTTCCTCCAAAT 1561
XX |||||||
201 GlyTyrTrpAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
XX |||||||
1562 GGGTACACAGCCGACACAGAGGGGCTTACGAGCTGTGCGACCTGTGAC 1611
XX |||||||
217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
XX |||||||
1612 CATCCGGGGCCCGAGGCCAGCTCGCCAGCTAATTAGCGCCCTGGCC 1661
XX |||||||
234 lnhIsArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
XX |||||||
1662 AGCACCGGAGAAACGATGTTGGAGAAATTCGTGGCTGATGGAAGAC 1711
XX |||||||
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
XX |||||||
1712 ACCACCCAGCTGGAACACTGACAACTACTCTCCGATGAGCCCAAGCCC 1761
XX |||||||
267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGlnAsnSera 284
XX |||||||
1762 GCTTAGCCCGAGCCCATCCAGCCCAAGCGAAACTTGAGATTCCG 1811
XX |||||||
284 lAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
XX |||||||
1812 CTCTCTGACGAGTGAGCCCTTCCACAGACAAACAAAGGCTTCTTC 1861
XX |||||||
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
XX |||||||
1862 GTGATGATGATGGAGACCCCTTCTCGCTGTGACTTAATCATCAGAGGCTC 1911
XX |||||||
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
XX |||||||
1912 CTCGCGCGCTGAGCAAGAGCGGTCTTATTATCAAAAGAAAGAGGACA 1961
XX |||||||
334 hrValleuArgGlnValAlaArgLeuAspProCysAspLeuGlnProIlePhe 350
XX |||||||
1962 CAGTGTGGCGGAGGTAGCTGAGCCCTGAGACCCCTGTGACTTCCAGCTATATCTTT 2011
XX |||||||

```

```
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGlu1 367
|||||
2012 GATGACATGCTCCACTTCTTAATCTGAGAGCTCGGGTATGATGAGA 2061
|||||
367 uileProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
|||||
2062 GATTCGCCCGCTGAGAGCAAACTAGACGGCTATTCGAAATTAATGGAG 2111
|||||
384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
|||||
2112 TCAGAGCCAGGAGACCCACCCACCTCTGACCTGTATTATACCAT 2161
|||||
401 LeuProAspLeuLeu 405
|||||
2162 CTTCCTGACCTCTG 2176

seq_name: /SIBS8/gcgdata/geneseq/geneseq/NA1999.DAT:AAV99927

seq_documentation_block:
ID AAV99927 standard; DNA; 3474 BP.
XX
AC AAV99927;
XX
XX 10-MAY-1999 (first entry)
XX
DE Novel human tumor necrosis factor receptor TR9.
XX
KW Human: tumor necrosis factor receptor; TNFR; TR9 receptor; cancer;
KW apoptosis; agonist; inhibition; autoimmune disorder; viral infection;
KW inflammation; antagonist; AIDS; neurodegenerative disorder; SS.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 247..2214
XX FT /tag= a
XX FT /product= "human TR9 receptor"
XX FT sig_peptide 247..366
XX FT /tag= b
XX FT mat_peptide 367..2211
XX FT /tag= c
XX
XX W09856892-A1.
XX
XX 17-DEC-1998.
XX
XX 10-JUN-1998; 98WO-US11932.
XX
XX 11-JUN-1997; 97US-0052991.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fan P, Gentz RL, Ni J, Yu G;
XX
XX WPI: 1999-060325/05.
XX
XX P-PSDB: AAW81059.
XX
XX New isolated tumor necrosis factor-like receptor, TR9 - used to
XX develop products for treating e.g. cancers, autoimmune disorders,
XX viral infections, inflammation, graft rejection, neurodegenerative
XX disorders or ischemic injury
XX
XX Claim 2; Fig 1; 134pp; English.
XX
XX This is the nucleotide sequence of the human tumour necrosis factor
XX receptor (TNFR), TR9 receptor, used in the method of the invention
XX to develop products to treat disorders such as cancers. The novel
XX TNFR, TR9, can be used to identify agents for modifying apoptosis.
XX Agonists can be used to treat diseases associated with increased
XX cell survival, or the inhibition of apoptosis, including cancers
XX (e.g. follicular lymphomas, carcinomas with p53 mutations, and
XX Kaposi's sarcoma and ovarian cancer), autoimmune disorders (e.g.
```

```
CC systemic lupus erythematosus and immune-related glomerulonephritis
CC rheumatoid arthritis), viral infections (e.g. herpes viruses, pox
CC viruses and adenoviruses), inflammation, graft vs host disease,
CC acute graft rejection and chronic graft rejection. Antagonists
CC can be used to treat diseases associated with increased apoptosis
CC including AIDS, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis,
CC retinitis pigmentosa, cerebellar degeneration), myelodysplastic
CC syndromes (e.g. aplastic anemia), ischemic injury (e.g. that
CC caused by myocardial infarction, stroke and reperfusion injury),
CC toxin-induced liver disease (e.g. that caused by alcohol), septic
CC shock, cachexia, anorexia, inflammatory diseases and stress response
CC related diseases, such as inflammatory bowel disease, rheumatoid
CC arthritis, osteoarthritis, psoriasis and septicemia. The products
CC can also be used for detection and diagnosis.
XX
XX Sequence 3474 BP; 832 A; 895 C; 844 G; 903 T; 0 other;

alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AAV99927 ..

Align seg 1/1 to: AAV99927 from: 1 to: 3474

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
997 ATGAACTCAACAGAAATCCAACTCTTCGCTCGTTAGACAAAGGTACT 1046
|||||
17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerAlaArgG 34
|||||
1047 GACTACATCCAGAGAGGAGGACAGCTCCCTGACAAACACAGCTCAGCAAGG 1096
|||||
34 LylsGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||
1097 GGAGGAAAGACGTGAACAAAGACCTCCAAACCTTCAGGTAGCAACAC 1146
|||||
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG 67
|||||
1147 CAGCAAGGCCCCACCACACACATCTGAGACTGCTCCGCTCCATGGA 1196
|||||
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
|||||
1197 GGCCACTGGGGGGAGAGAGTCCAGCAGGCCCATCAAGGGCCCAAGAGG 1246
|||||
84 LylHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
|||||
1247 GACATCTAGACAGACATACACAAAGCATTTTGACATGAGCATTTTG 1296
|||||
101 ProThrMetIleValLeuPheLeuLeuValLeuValIleValIleVal 117
|||||
1297 CCTGGATGATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1346
|||||
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
|||||
1347 GTCCAGTATCCGGAAGAAAGCTCGAGACTGTGAAAAAGGGGCCCGCAG 1396
|||||
134 sProSerAlaIleValIleLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1397 ATCCAGTGCATTTGTGAAAGAGGAGGCTGAAAGAAATTCAGATCTCCA 1446
|||||
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisLysIleAs 167
|||||
1447 ACCCAACACCGGGAGAAATGAGATCTACTGCAAAAGCCATGATGTCA 1496
|||||
167 PileLeuLysLeuValAlaAlaGlnValIleGlySerGlnTTPlysAspIleT 184
|||||
1497 TATCTGAAGCTTGTAGCAGCCCAAGTGGAGACCCAGTGGAAAGATATCT 1546
```

```

184 yrcGlnPheLeuCysAsnAlaSerGluValAlaAlaPheSerAsn 200
|||||
1547 ATCATGTTCTTTGCAATGCCAGTGAAGAGGAGTGTCTGTTCTCCAAT 1556
|||||
201 GLYTYrThrAlaAspHisGluArgAlaTYrAlaAlaLeuGlnHisTrpTh 217
|||||
1597 GGGTACACAGCCGACCGAGCGGCTGACGACCTGCGACGACTGAGAC 1646
|||||
217 rTLeaArgGlyProGluAlaSerLeuAlaGlnLeuLeuSerAlaLeuArg 234
|||||
1647 CATCCGGGGCCCGAGGCGAGCTCGCCAGCTAATTAGCGCCCTTGCGCC 1696
|||||
234 LnhIsArgAgaAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
|||||
1697 AGACCCGGAAGAAGCATGTTGTGGAGAAGATTCTGGCGCTGATGGAAGAC 1746
|||||
251 ThrThcGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1747 ACCACCCAGCTGGAACTGACAAACTAGCTCTCCGATGAGCCCGACCCC 1796
|||||
267 oluSerProSerProIleProSerProAlaAlaLysLeuGluAsnSerA 284
|||||
1797 GCTTAGCCCGAGCCCATCCCGCCAGCCCAAGCGAACTTGAGAATTCCG 1846
|||||
284 lAeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1847 CTCTCCGAGCGTGAGCCTTCCCGACAGACAAAGAACAGGCTTCTTC 1896
|||||
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
|||||
1897 GTGGATGATGTCGAGCCCTTCCTCCGCTGTGACTCTACATCCACGCGCTC 1946
|||||
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAsp 334
|||||
1947 CTCGCGCTGAGCAGCAAGCGTCTTTATTATACCAAGAAAGAGGACA 1996
|||||
334 hrValLeuAArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
1997 CAGTGTGGCGGCAAGTACGCCCTGAGCCCTGTGACTTGACGCCATATCTTT 2046
|||||
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluG 367
|||||
2047 GATGACACTGCTCCACTTTAAATCCCTGAGAGAGCTCGGGGTGATTGAAGA 2096
|||||
367 uileProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
|||||
2097 GATTCCCGCAGGCTGAGGACAACACTAGACCGGCTATTGGAATTAATGAG 2146
|||||
384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTYrSerHis 400
|||||
2147 TCAAGACCCAGGAAGCCAGACGACCTCTGAGCTCTGTTTATAGCCAT 2196
|||||
401 LeuProAspLeuLeu 405
|||||
2197 CTTCCTGACCTGCTG 2211
|||||
seq_name: /sids8/gcgdata/geneseq/geneseqn/MA2000.DAT:AAA99658
seq_documentation_block:
ID AAA99658 standard; cDNA; 3474 BP.
XX
AC AAA99658;
XX
XX 02-FEB-2001 (first entry)
XX
DE Human tumour necrosis factor receptor TR9 cDNA.
XX
XX Human tumour necrosis factor; TR9 receptor; immunosuppressive;
XX antiinflammatory; cardiant; antidiabetic; antiallergic;
XX antiallergic; antirheumatic; anti-HIV; anticonvulsant; cytostatic;
XX neuroprotective; gene therapy; Death domain containing Receptor 6;
XX common variable immunodeficiency; X-linked agammaglobulinemia;
XX severe combined immunodeficiency; Wiskott-Aldrich syndrome;

```

```

KW autoimmune disease; rheumatoid arthritis; allergic encephalomyelitis;
KW multiple sclerosis; diabetes mellitus; asthma; epilepsy; cancer;
KW cardiovascular disease; neurological disease;
KW protein coordinate data; ss.
XX
XX Homo sapiens.
FH
FT CDS location/Qualifiers
FT 247..2214
FT /*tag= a
FT /*product= "TR9 receptor"
FT sig_peptide 247..366
FT /*tag= b
FT mat_peptide 367..2214
FT /*tag= c
XX
XX WO200056862-A1.
XX
XX 28-SEP-2000.
XX
XX 16-MAR-2000; 2000WO-US06831.
XX
XX 24-MAR-1999; 99US-0126019.
XX
XX 14-MAY-1999; 99US-0134220.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX NI J, Gentz RL, Yu G, Fan P;
XX
XX WPI: 2000-594575/56.
XX
XX P-PSDB; AAB26981.
XX
XX Nucleic acid molecule encoding a human tumor necrosis factor receptor,
XX known as TR9, useful for treating, preventing or diagnosing severe
XX combined immunodeficiency, autoimmune diseases, HIV infection, epilepsy
XX and cancer -
XX
XX Claim 2: Fig 1: 220pp: English.
XX
XX PS The present sequence encodes a novel human tumour necrosis factor
XX receptor, designated TR9. The TR9 receptor is also known as Death Domain
XX containing Receptor 6. TR9 polypeptides, polynucleotides or agonists are
XX useful for treating, preventing or diagnosing common variable
XX immunodeficiency, X-linked agammaglobulinemia, severe combined
XX immunodeficiency and Wiskott-Aldrich syndrome, autoimmune diseases (such
XX as rheumatoid arthritis, allergic encephalomyelitis, multiple sclerosis,
XX diabetes mellitus and asthma), HIV infection, epilepsy, cancer,
XX cardiovascular diseases and other neurological diseases.
XX
XX SQ Sequence 3474 BP; 832 A; 895 C; 844 G; 903 T; 0 other;
XX
XX alignment_scores:
XX quality: 2091.00 length: 405
XX ratio: 5.163 gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-733-956-2 x AAA99658 ..
XX
XX Align seg 1/1 to: AAA99658 from: 1 to: 3474
XX
XX 1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
XX |||||||
XX 997 ATGACTCTACAGAAATCCAACTCTTGTGCTGTGTAGACCAAGTACT 1046
XX |||||||
XX 17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
XX |||||||
XX 1047 GAGTAGCATTCAGAGAGGACAGTCCCTGACACACAAAGCTCAGCAAGG 1096
XX |||||||
XX 34 LylsGluAspValAlaAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
XX |||||||
XX 1097 GGAAGGAAGAGTGAACAGACCTCCCAACCTTCAGGTAGTCAACCAC 1146
XX |||||||

```

51 GInGInGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG1 67
 |||||
 1147 CAGCAAGGCCCCACCACAGACACATCTGAAGCTGCTGCCGTCATGGA 1196
 67 uAlaThrGlyGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
 |||||
 1197 GGCACCTGGGGGCGAAGATCCAGCACGCCCATCAAGGGCCCCCAAGAGG 1246
 84 LHisProArgGlnAsnLeuHisLysHisPheSerPheLysngLHisLeu 100
 |||||
 1247 GACATCCTAGACAGAACTACACAGCATTTTGACATCATAGCATTTGG 1296
 101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValIleVal 117
 |||||
 1297 CCTGGATGATTTGTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1346
 117 LysSerIleArgLysSerSerArgThrIleuLysLysGlyProArgGln 134
 |||||
 1347 GTGCAGTATCCGGAAGAGCTCGAGGACTGTGAAAAAGGGGGCCCCGAG 1396
 134 sPProSerAlaIleValGlyLysAlaGlyLeuLysLysSerMetThrPro 150
 |||||
 1397 ATCCCAAGTGCATTTGTGAAAGGGCGAGGCTGAGAAATCCATGACTCCA 1446
 151 ThrGlnAsnArgGlyLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
 |||||
 1447 ACCCAGAACCCGGAGAAATGATCTACTGCAATGGCCATGTGTGCA 1496
 167 PLeuLeuLysLeuValAlaAlaGlnValLysSerGlnTrpLysAspIleT 184
 |||||
 1497 TATCCCGAAGCTTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1546
 184 YrGlnPheLeuLysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
 |||||
 1547 ATCAGTTTCTTTCGCAATGCCAGTGAAGAGGAGTTGCTGCTTCTCCAA 1596
 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
 |||||
 1597 GGGTACACAGCGACACGAGCGGGGCTACGACGCTCAGACACTGAGAC 1646
 217 rIleAsnGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
 |||||
 1647 CATCCGGGGCCCCGAGCCAGCTCCGCAATTAATTACGCCCTCTGCC 1696
 234 LHisArgArgAsnAspValValGlyLysIleArgGlyLeuMetGluAsp 250
 |||||
 1697 ACGACCGGAGAAACGATGTTGTGGACAAAGATTCGTGGCTGATGAGAAC 1746
 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
 |||||
 1747 ACCACCCACTGGAAACTGACAAACTAGCTCTCCGATGAGCCCCAGCCC 1796
 267 OleuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
 |||||
 1797 GCTTACCCCGAGCCCATATCCCAAGCCCCCAACGGAACTTGAGAAATTC 1846
 284 lAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
 |||||
 1847 CTCTCCTGAGGTGAGGCTTCCGCCACAGACAAAGAGGCTTCTTTC 1896
 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerLysSe 317
 |||||
 1897 GTGGATGAGTCCGAGCCCTTCTCCCTGTGACTTACATCCAGCGGCTC 1946
 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
 |||||
 1947 CTCGGCGCTGAGCAGAGACGGTTCCTTTATTTACCAAGAAAGAAAGACA 1996
 334 hValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
 |||||
 1997 CAGTGTGGGCGAGTACGCTGAGACCCCTGTGACTTGCAGCCTATCTTT 2046

351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluC1 367
 |||||
 2047 GATGACATGCTCCACTTCTTAATCTCTGAGAGCTCGGGTGATTGAGA 2096
 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleGlyLV 384
 |||||
 2097 GATTCGCCAGGCTGAGAGCAAACTAGACCGGCTATTGAAATTAATGGAG 2146
 384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
 |||||
 2147 TCAGAGGCCAGAGAGCCAGCCAGACCCCTCCTGAGCTGTGATTATACCAT 2196
 401 LeuProAspLeuLeu 405
 |||||
 2197 CTTCCTGACCTGCTG 2211
 seq_name: /SIDS8/gcdata/geneseq/geneseq/M1199.DAT:AA233945
 seq_documentation_block:
 ID AA233945 standard; cDNA; 3534 BP.
 AC AA233945;
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO868 nucleotide sequence.
 XX
 KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KW secreted protein; transmembrane protein; ss.
 OS Homo sapiens.
 XX
 PN W09946281-A2.
 XX
 PD 16-SEP-1999.
 XX
 EF 08-MAR-1999; 99W0-US05028.
 XX
 PR 10-MAR-1998; 98US-0077450.
 PR 11-MAR-1998; 98US-0077632.
 PR 11-MAR-1998; 98US-0077641.
 PR 11-MAR-1998; 98US-0077649.
 PR 12-MAR-1998; 98US-0077791.
 PR 13-MAR-1998; 98US-0078004.
 PR 17-MAR-1998; 98US-0040220.
 PR 20-MAR-1998; 98US-0078886.
 PR 20-MAR-1998; 98US-0078910.
 PR 20-MAR-1998; 98US-0078936.
 PR 20-MAR-1998; 98US-0078939.
 PR 25-MAR-1998; 98US-0079294.
 PR 26-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079663.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 27-MAR-1998; 98US-0079728.
 PR 27-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 30-MAR-1998; 98US-0079923.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 31-MAR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 01-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081071.
 PR 08-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 09-APR-1998; 98US-0081229.

PR	15-APR-1998:	98US-0081817.
PR	15-APR-1998:	98US-0081838.
PR	15-APR-1998:	98US-0081852.
PR	15-APR-1998:	98US-0081955.
PR	21-APR-1998:	98US-0082568.
PR	21-APR-1998:	98US-0082569.
PR	22-APR-1998:	98US-0082700.
PR	22-APR-1998:	98US-0082704.
PR	22-APR-1998:	98US-0082804.
PR	23-APR-1998:	98US-0082767.
PR	27-APR-1998:	98US-0082796.
PR	28-APR-1998:	98US-0083336.
PR	28-APR-1998:	98US-0083322.
PR	29-APR-1998:	98US-0083392.
PR	29-APR-1998:	98US-0083495.
PR	29-APR-1998:	98US-0083496.
PR	29-APR-1998:	98US-0083499.
PR	29-APR-1998:	98US-0083500.
PR	29-APR-1998:	98US-0083545.
PR	29-APR-1998:	98US-0083554.
PR	29-APR-1998:	98US-0083558.
PR	29-APR-1998:	98US-0083559.
PR	30-APR-1998:	98US-0083742.
PR	05-MAY-1998:	98US-0084366.
PR	06-MAY-1998:	98US-0084414.
PR	06-MAY-1998:	98US-0084441.
PR	07-MAY-1998:	98US-0084598.
PR	07-MAY-1998:	98US-0084600.
PR	07-MAY-1998:	98US-0084627.
PR	07-MAY-1998:	98US-0084637.
PR	07-MAY-1998:	98US-0084639.
PR	07-MAY-1998:	98US-0084640.
PR	07-MAY-1998:	98US-0084643.
PR	13-MAY-1998:	98US-0085323.
PR	13-MAY-1998:	98US-0085338.
PR	13-MAY-1998:	98US-0085339.
PR	15-MAY-1998:	98US-0085573.
PR	15-MAY-1998:	98US-0085579.
PR	15-MAY-1998:	98US-0085580.
PR	15-MAY-1998:	98US-0085582.
PR	15-MAY-1998:	98US-0085589.
PR	15-MAY-1998:	98US-0085697.
PR	15-MAY-1998:	98US-0085700.
PR	18-MAY-1998:	98US-0086023.
PR	22-MAY-1998:	98US-0086392.
PR	22-MAY-1998:	98US-0086414.
PR	22-MAY-1998:	98US-0086430.
PR	22-MAY-1998:	98US-0086486.
PR	28-MAY-1998:	98US-0087098.
PR	28-MAY-1998:	98US-0087106.
PR	28-MAY-1998:	98US-0087208.
PR	30-JUL-1998:	98US-0094651.
PR	11-SEP-1998:	98US-0100038.

PA (GETH) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
XX WPI; 1999-551358/46.
DR
DR P-PSDB; AA141693.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
XX
XX
PS Claim 2; Fig 25; 530P; English.
XX

CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AA233891 to
CC AA234318, and AA411685 to AA411774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.

XX
XX Sequence 3534 BP, 829 A; 927 C; 958 G; 920 T; 0 other;

```
alignment_scores:
  quality: 2091.00      Length: 405
  ratio: 5.163          Gaps: 0
  percent similarity: 100.000  Percent identity: 100.00000
```

```
alignment_block:
US-09-733-956-2 x AAZ33945 .
```

Align seg 1/1 to: AAZ33945 from: 1 to: 3534

1 MetAspErThrGluSerAsnSerSerLaaSerValaArgPolySValle 17
1075 ATGAATCTCAACAGAAATCCAACTCTTCTCCTCTGTTAAACCAAAAGGACT 1124
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerLaaArg 34
1125 GAGTACGATCCAGGAAGGACAGTCCCTGACACAAACGACTCAGCCAGG 1174
34 LYLysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
1175 GGAAGGAAGACGTGAACAGAGACCCCTCCAAACCTTCAAGTATGATCAACAC 1224
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG1 67
1225 CAGCAAGGCCCCCACCACAGACACTCTGAGAGCTGCTGCCGCTCCATGGA 1274
67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 84
1275 GGGCATTGGGGGCGAAGATGCCAGCAGCCCATCAAGGCGCCCAAGAGG 1324
84 LHisProArgGlnAsnLeuHisLysHisPheAspP1easGluHisLeu 100
1325 GACATCTTGACAGAACCTACACAAAGCATTTTGACATCAAGACCATTTG 1374
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleVala 117
1375 CCTGGATATTTGTCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1424
117 LcysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlna 134
1425 GTGCATATTCGGAAAGCTCGAAGACTGTGAAAAAGGGGCGCCGACAG 1474
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
1475 ATCCCACTGCCATTTGTGAAAAAGGCGAGGGCTGAGAAATTCATGACTCCA 1524
151 ThrGlnAsnArgGluLysTrpIleLeuThrLysAsnGlyHisGlyIleAs 167
1525 ACCCAAGACCGGGAGAAATGATCTACTACTGCAATGGCCATGGATGTGA 1574
167 P1eLeuLysLeuValAlaAlaGlnValaG1SerGlnTrpLysAspIleT 184
1575 TATCCTGAAGCTTGTAGCAGGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1624
184 yrcGlnPheLeuLysAsnAlaSerGluThrGluValAlaAlaPheSerAsn 200
1625 ATCAGATTTCTTGCAATGCCACTGAGAGGGAGGTGTGCTTCTCCCAT 1674
201 G1ATYTrpThAlaAspHisGlnArgLarTrpAlaAlaLeuGlnHisTrpTh 217
1675 GGGTACACAGCCGACACGAGGGGGGCTTACGACGAGCTTGCACGACTGGAC 1724
217 r1leArGcGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234

```

|||||
1725 CATCCGGGGCCCCGAGGCCAGCTCCGCACTAATTAGCGCCCTCGCC 1774
234 |nHisrgrgrsnaSpvalValglulysllearglleuMetgluAsp 250
1775 AGCAGCGGAGAAACGATGTTGTGAGAAAGTTCGTGGCGTGATGAGAAC 1824
251 |ThrThgluGlulThrAspLysleuAlaLeuProMetSerProSerP 267
1825 ACCACCCAGCTGGAACTGACAACTAGCTCTCCGATGAGCCCCAGCCC 1874
267 |leuSerProSerProIleProSerProAsnAlaLysleuGlusSerA 284
1875 GCTTACCCGAGGCCCATCCGCCAGCCCCAACGCAACTGAGAAATTCG 1924
284 |leuLeuThrValGluproSerProGlinspLysAsnLysgliphep 300
1925 CTCTCTGACGGGGAGCCTTCCCCACAGACAGAACAGAGGCTTCTTC 1974
301 |ValAspGlusergluProleuLeuArgCysAspSerThrSerSergly 317
1975 GTGGAGAGTCCGGAGCCCTTCTCCGCTGTGACTCATCATCCAGCGGCT 2024
317 |fSerAlaLeuSerArgAsnGlySerPheIleThrLysglulysLysAsp 334
2025 CTCCGGCTGAGCAGACGAGGTTCTTTATACCAAGAAAGAGAGAGACA 2074
334 |hValleuArggluValArgleuAspProCysAspLeuGluproIleph 350
2075 CAGTGTGGCGGAGTACGCTGAGCCCTGTCGACTTGACAGCTATCTTT 2124
351 |AspAspMetleuHisPheleuAsnProgluGlulLeuArgValIlegl 367
2125 GAGGACATCTCCACTTCTTAATCTGAGAGCTCGGGGTGATTGAAGA 2174
367 |uileProgluAlaGlusPlysleuAspArgleuPhegluIleIleGly 384
2175 GATTCGCCAGGCTGAGACAACTAGACCCGCTATTGCAATATTGGAC 2224
384 |AllySerGlnGluAlaSerGlnThrleuLeuAspSerValTyrSerHis 400
2225 TCAGAGGCCAGGAGGACGACGACCTCTGAGCTGTGTTATAGCCAT 2274
401 |leuProAspLeuLeu 405
2275 CTTCCTGACCTCTGCTG 2289

seq_name: /SID58/gcdata/geneseq/geneseqn/NA2000.DAT: AAC78474
seq_documentation_block:
ID AAC78474 standard; cDNA: 3534 BP.
XX AAC78474;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human PRO868 (UN0437) nucleotide sequence SEQ ID NO:63.
XX
XX Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
XX expressed sequence tag; detection; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200053756-A2.
XX
XX 14-SEP-2000.
XX
XX 18-FEB-2000; 2000WO-US04341.
XX
XX 08-MAR-1999; 99WO-US05028.
XX 12-MAR-1999; 99US-0123957.
XX 29-MAR-1999; 99US-0126773.
XX 21-APR-1999; 99US-0130232.

```

```

PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Destroyers L, Eaton DL;
XX Ferrara N, Filvarolf E, Fong S, Gao W, Gerber H, Gerltsen ME;
XX Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
XX Kljavin ID, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
XX Shelton DL, Stewart RA, Tumas D, Williams PM, Wood WI;
XX
XX WPI: 2000-611443/58.
XX P-PSDB: AAB44249.
XX
XX Novel PRO polypeptides and polynucleotides used in detection methods,
XX to target bioactive molecules to specific cells, and to modulate
XX cellular activities -
XX
XX Claim 2; Fig 25; 636pp; English.
XX
XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed
XX sequence tag) sequences which encode secreted or transmembrane PRO
XX polypeptides. The PRO polynucleotides and polypeptides have cytosolic
XX activity. The polynucleotides and polypeptides can be used for detecting
XX the presence of PRO polypeptides in samples, for linking bioactive
XX molecules to cells and for modulating biological activities of cells,
XX using the polypeptides for specific targeting. The polypeptide targeting
XX can be used to kill the target cells, e.g. for the treatment of cancers.
XX The polypeptide pairs provide specific targeting of bioactive molecules
XX to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
XX the isolation of the PRO polynucleotide sequences.
XX
XX SQ Sequence 3534 BP; 829 A; 927 C; 858 G; 920 T; 0 other.

alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AAC78474 ..
Align seg 1/1 to: AAC78474 from: 1 to: 3534

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
|||||
1075 ATGAACTCAACAGAAATCCAACTCTTGCGCTGTGAGACCAAGTACT 1124
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
1125 GAGTACATCCACAGGAGGAGACACTCCCTGACAAACACACAGCTCAGCAAGG 1174
34 |lyysgluAspValasnlusThrleuProAsnleuGlnValValasnhis 50
1175 GGAAGGAGAGCGTGAACAGAGACCTCCCAACCTTGAGGTAGCAACCCAC 1224
51 |GlnGlnGlyProHisHisArgHisIleleuLysleuLeuProSerMetgl 67
|||||
1225 CAGCAAGGCCGCCACCAAGACACATCTCGAAGCTGCTCGCTCATGGA 1274

```



```
67 uAlaThrGlyGlyLysSerSerThrProLeuGlyProLysArg 84
|||||
1275 GGGCAGTGGGGGAGAGAGTCCAGCAGCCCATCAAGGCCCAAGAGG 1324
84 LysProArgGlnAsnLeuHisLysHisPheAspLeuGlnHisLeu 100
|||||
1325 GACATCTTAGAGAACTACACAAAGCATTTTGACATCAATGACATTTG 1374
101 ProTrpMetLeuValLeuPheLeuLeuValLeuValValLeuVal 117
|||||
1375 CCTGGATGATTTGGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1424
117 LysSerLeuArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
|||||
1425 GTGCAGATCTCGGAAAAGCTCGAGAGACTGTGAAAAGGGGGCCGCAAG 1474
134 sPProSerAlaIleValGlyLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1475 ATCCCACTGCATTTGTGAAAAGGAGGAGGCTGAGAAATCATGACTGCA 1524
151 ThrGlnAsnArgGlyLysThrPheTyrTyrCysAsnGlyHisGlyIleAs 167
|||||
1525 ACCAGAGACCGGGGAAATGATCTACTGCAATGCGCATGCTATCGA 1574
167 PileLeuLysLeuValAlaIleValGlySerGlnTrpLysAspIleT 184
|||||
1575 TATCCTCAACTTGTAGCAGCCCAAGTGGGAGGCCAGTGGAAAGATATCT 1624
184 TrpGlnPheLeuCysAsnAlaSerGlnArgGlnValAlaAlaPheSerAsn 200
|||||
1625 ATCAGTTCTTTGCAATGCAGTGCAGAGGAGGCTGCTGCTTCTCCAAAT 1674
201 GlyTyrThrIleAspHisGlyArgGlyAlaTyrAlaAlaLeuGlnHisTrpTh 217
|||||
1675 GGGAGACACCGCCGACGAGCGGGGCTACGACCTCTGCGACACTGGAC 1724
217 rIleArgGlyProGlnLaserLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||
1725 CATCGGGGGCCGAGCGCCAGCTCGCCAGCTAATTAAGCGCCCTGGGCC 1774
234 LnhIsArgArgAsnAspValValGlyLysIleArgGlyLeuMetGlnAsp 250
|||||
1775 AGCACCAGAGAAACGATGTTGTGAGAAAGATTCGTGGCTGATGAAAGAC 1824
251 ThrThrGlnLeuGlnHisAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1825 ACCACCCAGCTGGAAGCTGAAGAACTGACTCTCCGATGAGCCCCACGCC 1874
267 OLeuSerProSerProIleProSerProAsnAlaLysLeuGlnAsnSera 284
|||||
1875 GCTTACGCCGAGCCCATCCCAAGCCCAACGCAAACTTGAGAAATTCGC 1924
284 LaleuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1925 CTCCTCCAGAGGTGGAGCTTCCACAGCAGCAAGAACAGGCTTCTTTC 1974
301 ValAspLysLeuSerGluProLeuLeuArgCysAspSerThrSerGlySe 317
|||||
1975 GTGGATAGTGGAGGCCCTTCTCCGTGTGACTCTGACTCATCCACGCGCTC 2024
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLysLysAspT 334
|||||
2025 CTCGCGCCTGAGCAGAACGGTTCCTTATTATACAAAGAAAAGAGACA 2074
334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
2075 CAGGTGTCGGCAGGTACGCTGGAGCCCTGTGACTTGACACTATCTT 2124
351 AspAspMetLeuHisPheLeuAsnProGlnGluLeuArgValIleGluG 367
|||||
2125 GATGACATGTGCTCCACTTCTTAATCTGAGAGAGCTGGCGGATTTGAAGA 2174
367 uIleProGlnAlaGlnLysLysLeuAspArgLeuPheGlnIleIleGly 384
```

```
|||||
2175 GATTCGCCAGGCTGAGGACAAACTAGACCGGCTATTCGAATTAATTGGAG 2224
384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
|||||
2225 TCAAGAGCCAGGAGGAGCCAGCAGACCTCTGACTCTGTTATATAGCAT 2274
401 LeuProAspLeuLeu 405
|||||
2275 CTTCCTGACCTGCTG 2289

seq_name: /SID8/gcgcdata/geneseq/geneseqn/NA2001.DAT:AAH14374
seq_documentation_block:
ID AAH14374 standard; cDNA: 2636 BP.
AAH14374;
26-JUN-2001 (first entry)
DE Human cDNA sequence SEQ ID NO:11786.
XX Human: primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX MPI: 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 11786; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination
XX of the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
```

CC of the present invention.
 XX Sequence 2636 BP; 569 A; 762 C; 722 G; 583 T; 0 other;
 SQ

alignment_scores:
 Quality: 2087.00 Length: 405
 Ratio: 5.153 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.753

alignment_block:
 US-09-733-956-2 x AAH14374

Align seg 1/1 to: AAH14374 from: 1 to: 2636

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
  |||||||
1145 ATGAATCTACACGAATCCAACTCTTCTGCTCTGTAGACCAAGGTA 1194
  |||||||
17 userSer11eGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
  |||||||
1195 GAGTAGCATCCAGGAAGGACAGTCCCTGACAACACAGCTCAGCAAGG 1244
  |||||||
34 LylsGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
  |||||||
1245 GGAAGGAAAGACGTGAACAGACCTCCCAACCTTCAGGTAGTCAACCC 1294
  |||||||
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG 67
  |||||||
1295 CAGCAAGGCCCCCACCACAGACATCTCGAAGCTGCTCCCTCCATGSA 1344
  |||||||
67 uAlaThrGlyGlyGluLysSerSerThrProLysGlyProLysArgG 84
  |||||||
1345 GGCACATGGGGCGAGAAAGTCCAGCACGCCCATCAAGGGCCCCAAGAG 1394
  |||||||
84 LylHisProArgGlnAsnLeuHisLysHisPheAspTleAsnGlnHis 100
  |||||||
1395 GACATCTCAAGACAGAACTACACAGCAATTTGACATCAATGAGCATTT 1444
  |||||||
101 ProTyrMetIleValLeuPheLeuLeuValIleValIleValIle 117
  |||||||
1445 CCTGGATGATTTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
  |||||||
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
  |||||||
1495 GTGCAATATCCGGAAGAAAGCTCGAGCACTCTGAAGAGGGCCCCGAG 1544
  |||||||
134 sPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetPro 150
  |||||||
1545 ATCCCAAGTGCATTTGTGAAAGAGGAGGCTGAGAAATCCATGACTCCA 1594
  |||||||
151 ThrGlnAsnArgGluLysTyrIleTyrTyrCysAsnGlyHisGlyTle 167
  |||||||
1595 ACCCAAGAACCGGGAAGAAATGATCTACTACTGCAATGGCCATGATGCA 1644
  |||||||
167 PileuLeuLysLeuValAlaGlnValGlySerGlnTyrPylsAspIle 184
  |||||||
1645 TATCCGGAAGCTTTGACAGCCCAAGTGGAAAGCCAGTGGAAAGATATCT 1694
  |||||||
184 YrGlnPheLeuCysAsnAlaSerGluArgGluValAlaIlePheSerAsn 200
  |||||||
1695 ATCAGATTCTTTGCAATGCGAGAGAGGAGGTTCCTGCTTCTCCAAAT 1744
  |||||||
201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaIleLeuGlnHisTyr 217
  |||||||
1745 GGGTAGACACGCGACCAAGAGGCGGCTACGCACTCTGACAGACTGGAC 1794
  |||||||
217 r11eArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
  |||||||
1795 CATCCGGGGCCCGGAGGCGAGCTCGCCAGTAATTAGGCGCCCTGGCC 1844
  |||||||
234 LnlHisArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
  |||||||

```

```

1845 ACCACCGGAAGAAAGATGTTGTGAGAAAGATTCTGGGCTGATGGAAGAC 1894
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
  |||||||
1895 ACCACCGACTGGAAAGTACAACTAGCTCTCCGATAGGCCCAAGCCC 1944
  |||||||
267 OleuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSera 284
  |||||||
1945 GCTTAGCCCGACCCCATCCACGCCCAACCGAAACTTGAGAAATTCGG 1994
  |||||||
284 lAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
  |||||||
1995 CTCTCCTGACGGTGGAGCCTTCCCAACAGACACAGAGGCTTCTTC 2044
  |||||||
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
  |||||||
2045 GTGATGATGATCGGAGGCCCTTCTCGCTGTGACTGATCATCCAGCGGCT 2094
  |||||||
317 fSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLysLysAspT 334
  |||||||
2095 CTCCGCGCTGAGCAGGACGATTCCTTTATTACAAAGAAAGAAAGAGACA 2144
  |||||||
334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
  |||||||
2145 CAGTGTTCGCGAGAGTACGCTGGACCCCTGTGACTTGCAGCCATCTTT 2194
  |||||||
351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluG 367
  |||||||
2195 GATGACATGCTCCACTTCTTAATCTCGAGAGCTCGGGGTATTGAGACA 2244
  |||||||
367 u11eProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGly 384
  |||||||
2245 GATTCCCCAGCGCTGAGAGACAACTAGACCGGCTATTGCAATTTATGGAG 2294
  |||||||
384 allYserGlnGlnAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
  |||||||
2295 TCAAGAGCCAGAGGACGACCAACCTTCCTGGACTCTGTTATAGCCAT 2344
  |||||||
401 LeuProAspLeuLeu 405
  |||||||
2345 CTTCCTGACCTGCTG 2359
  |||||||
seq_name: /SIDS8/ycgdata/geneseq/geneseqn/NA1999.DAT:AAx25320
seq_documentation_block:
ID AAX25320 standard; cDNA: 3331 BP.
XX
AC AAX25320:
XX
DT 19-JUL-1999 (first entry)
XX
DE Mouse TNF receptor TRL cDNA.
XX
KW TRL; TNF receptor-like; tumour necrosis factor receptor; mouse;
KW signal transduction; cell differentiation; prostate cancer;
KW inflammation; arthritis; diabetes; insulin resistance; diagnosis;
KW therapy; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 344..2065
FT FT /*tag= a
XX
XX MO9915663-A1.
XX
PD 01-APR-1999.
XX
XX 25-SEP-1998; 98WO-US20219.
XX
XX PF 17-MAR-1998; 98US-0042785.
XX
XX PR 26-SEP-1997; 97US-0938896.
XX

```



```
seq_name: /SID8/gcdata/geneseq/geneseqn/NA199. DAT: AAX25264
2045 CATCTTCTGACCTATTG 2062
|||||
seq_documentation_block:
ID AAX25264 standard; DNA: 1893 BP.
XX AAX25264;
AC
XX
XX
XX 19-JUL-1999 (first entry)
DE Degenerate nucleotide sequence encoding ZTNFR-6.
XX
XX ZTNFR-6; tumour necrosis factor receptor-6; human;
XX cell maturation; bone cell regulation; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9111790-A1.
XX
XX 11-MAR-1999.
XX
XX 03-SEP-1998; 98WO-US18364.
XX
XX 04-SEP-1997; 97US-0923725.
XX 04-SEP-1997; 97US-0057608.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Farrah TM, Gross JA, Matthews SM;
XX
XX WPI: 1999-205190/17.
XX
XX P-PSDB: AAY05678.
XX
XX New secreted or membrane bound tumor necrosis factor receptor
XX ZTNFR-6 - useful for detecting a genetic abnormality in a patient
XX
XX Claim 10c; Page 129-130; 145pp; English.
XX
XX This degenerate nucleotide sequence encodes a novel human
XX membrane-bound tumor necrosis factor receptor, designated ZTNFR-6
XX (see AAY05678). It includes all possible codons for the given
XX amino acid sequence of ZTNFR-6. The degenerate sequence can
XX serve as a template for optimising expression of ZTNFR-6
XX polynucleotides in various cell types and species. Sequences
XX containing preferential codons can be tested and optimised for
XX expression in various species, and tested for functionality.
XX ZTNFR-6 is a novel membrane of the TNFR family. ZTNFR-6
XX polypeptides can be used to promote cellular maturation and
XX bone cell regulation. Polynucleotides can be used in the
XX recombinant production of ZTNFR-6 polypeptides, and in a method for
XX detecting a genetic abnormality in a patient.
XX
XX Sequence 1893 BP; 338 A; 224 C; 293 G; 211 T; 827 other;
SQ
alignment_scores:
Quality: 1567.00 Length: 403
Ratio: 4.650 Gaps: 0
Percent Similarity: 83.623 Percent Identity: 73.449
alignment_block:
US-09-733-956-2 x AAX25264 ..
Align seg 1/1 to: AAX25264 from: 1 to: 1893
1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProIysValIle 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
679 ATGAAYWSNACNGARWMSNAAYWSMWSNGCWNMGNTNGMCNNAARCTNYT 728
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
```

```
.....|
729 NWSNWSNATHCARGARGNACNGTNCNGAYAAACNWSNWSNGCWMGNG 778
34 lylYsglAspValAsnLysThrLeuProAsnLeuIleValIleAsnHis 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
779 GNAARGARGAYGTNAAYAAACAYTNCNNAAYTTNCARGTNGTNAAYCAY 828
51 GlnGlnGlyProHisIleArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
829 CARGARGNCCNCAAYCAYMGNCAYATHYTNAARYTYTNCNMSNATGGA 878
67 uAlaThrGlyGlyLysSerSerThrProIleLysGlyProIysArg 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
879 RGCNACNGNGNGARGARARWMSNMACNCCNATHARGGNCNAARMGNG 928
84 lylHisProArgGlnAsnLeuHisLysHisIleAspIleAsnGlnHisLeu 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
929 GNCAYCCNMGNCARAAAYTTNCATTAARCAATTTTAAAYGARGAYTN 978
101 ProIleMetIleValLeuPheLeuLeuValLeuValIleValIle 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
979 CWNMGATGATGCTGCTNTTYTNTTYTNGTNGTNGTNGTNGTNGTNGT 1028
117 lGysSerIleArgLysSerSerArgThrLeuLysGlyProArgGlnA 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1029 NTGYWSNATHMGNAARWMSNMGNCNNTNAARARGGNCNMGNCARG 1078
134 sProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1079 AYCCNMSNCGNATHGTNGARARGCNGNTNAARARWMSNATGACNCCN 1128
151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1129 ACNCAARAAYWNGARARARAGTGAATHTAATGYAAAGCNCAYGCGNATGCA 1178
167 PileLeuLysLeuValAlaIleGlnValGlySerGlnTrpLysAspIle 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1179 YATHYTNAARYTNGTNGCNCARGTNGSNCARGTGAAGATATHT 1228
184 yrlGlnPheLeuCysAsnAlaSerGluArgGluValAlaIleAspSerAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1229 AYCARTTYTNTGYAAYGCMWSNGARWNGARGTNGCNGTNTYWSNAY 1278
201 GlyTyrThrIleAspHisGluArgAlaTyrAlaIleLeuGlnHisIleTrp 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1279 GCGTAYACNCCNGAYCAYGARMGNGCNTAYGCGNCTNCARAYTGAC 1328
217 rIleArgGlyProGlnAlaSerLeuAlaGlnIleLeuIleSerAlaLeuArg 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1329 NATHMGNGCNCARGARCNWSNTTNGCNCARYTNATHWSNGCTYTMGNC 1378
234 lnhIsArgArgAsnAspValAlaGluLysIleArgGlyLeuMetGluAsp 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1379 ARCAAYWGMNGMAAYGAVGTNGTNGARARARATHMGNGNTYTNATGARGAY 1428
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1429 ACNACNCAARYTNGARACNAGAYTYTNGCYTNCNATGMSNCCNMSNCC 1478
267 oluSerProSerProIleProSerProAsnAlaLysLeuGlnAsnSerA 284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1479 NYTNWSNCCNWSNCCNATHCCNMSNCCNMAAYGCMMAARYTNGAATAAAYWSN 1528
284 lAleuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1529 CNTYTNACNCTNGARGCNCNWSNCCNARGAYARAARAYARGNTTYTY 1578
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1579 GTNGAYGARGWSNARGCNCNTYTNMGNTGYGAYWSNCCNMSNWSN 1628
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysAspTr 334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

CC	seq_name:	/SIO58/gcgcdata/geneseq/geneseqn/NA1999.DAT:AAx25322
XX	seq_documentation_block:	
XX	ID	AAx25322 standard; cDNA: 2638 BP.
XX	AC	AAx25322;
XX	DT	19-JUL-1999 (first entry)
XX	DE	Human TNF receptor TRL II CDNA.
KX	TRL II: TNF receptor-like; tumour necrosis factor receptor; human;	
KW	signal transduction; cell differentiation; prostate cancer;	
KW	Inflammation; arthritis; diabetes; insulin resistance; diagnosis;	
XX	therapy; ss.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	510..2227
FT	s1g_peptide	/*tag= a 510..632
FT	mat_peptide	/*tag= b 633..2324
XX	WO9915663-A1.	
XX	01-APR-1999.	
XX	25-SEP-1998;	98WO-US20219.
XX	17-MAR-1998;	98US-0042785.
PR	26-SEP-1997;	97US-0638896.
PA	(MILL-) MILLENNIUM BIOHERAPEUTICS INC.	
XX	Busfield SJ;	
DR	WPI: 1999-254712/21.	
DR	P-PSDB: AAY05697.	
PT	Nucleic acid encoding tumor necrosis factor receptor-like protein	
XX	Claim 1b: Page 148-153; 170pp; English.	
XX	This cDNA sequence codes for a novel member of the human tumour	
CC	necrosis factor receptor superfamily (see AAY05697), termed TRL II	
CC	(TNF receptor-like). A plasmid containing the full-length	
CC	nucleotide sequence encoding human TRL II (clone ephT51) is	
CC	deposited as ATCC 98649). TRL II has a mol.wt. of 66.2 kDa.	
CC	It is a membrane-bound protein that has 4 cysteine-rich domains	

CC and a C-terminal death domain (see AAY05698). TRL II cDNA was
CC isolated from a HeLa cDNA library using mouse TRL cDNA (see
CC AA535320) as probe. An alternatively spliced, secreted form, human
CC TRL I (see AAX55321), has also been identified. Chromosomal mapping
CC located the TRL gene on chromosomal 6p25. The invention provides
CC full-length TRL proteins, TRL fusion proteins, antigenic peptides
CC and anti-TRL antibodies, as well as TRL nucleic acids, recombinant
CC expression vectors, host cells and non-human transgenic animals.
CC TRL are regulators or modulators of cellular signal transduction,
CC cellular proliferation or differentiation, cell survival and
CC apoptosis, immune system cells, and cells involved in insulin
CC resistance or the diabetic response. They can be used to isolate
CC cognate ligands, to modulate ligand/TRL interactions, to screen
CC for potential modulators, and to treat conditions associated with
CC abnormal TRL levels. TRL nucleic acids are used for recombinant
CC production of TRL, to detect mutations, to identify TRL homologues,
CC as antisense, triplex-forming or ribozyme therapeutics, for
CC chromosome mapping and for forensic identification of individuals.
CC Antibodies, antisense sequences, TRL protein mutants etc. that
CC modulate activity of TRL are used to treat or prevent conditions
CC associated with aberrant TRL protein or expression, e.g. cell
CC proliferation and differentiation disorders (particularly cancer,
CC especially of the prostate), inflammation (e.g. arthritis) and
CC metabolic disease (e.g. diabetes and insulin resistance).
XX
XX

```
alignment_scores:      Length: 294  
    quality: 1527.00     Gaps: 0  
    Ratio: 5.194  
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-733-956-2 X AAX25322

Align seg 1/1 to: AAX25322 from: 1 to: 2638

1	MetaLsSerThnCluSerAmsSerSerAlaSerValArgProIlysVal	17
1260	ATGAACCTCAACAGAAATTCACACCTTCTCTCTCTGTGTAAACCAAGAGTACT	1309
17	userSerIleGlnGlnIuGlyThValProAspAsnThrSerSerAlaArg	34
1310	GAGTACACATCCAGGAAGGAGCAGTCCCTGACACACCAAGCTCAGCAAGG	1359
34	LYLysGluAspValAsnIysThrLeuProAsnLeuGlnValAlaAsnHis	50
1360	GGAGAGAAACCTCGAACAAGACCTCCCAAACTTCACAGTGTCAACACAC	1409
51	GlnGlnGlyProHisIleThrIleLeuIysLeuLeuProSerMetG1	67
1410	CAGCAAGGCCCCCACCACACAGACACATCTGAAAGCTGCTCCGTCATGGA	1459
67	uAlaThrGlyGlyGluIuLysSerSerThrProIleLysGlyProLysArg	84
1460	GGCCACATGGGGCGGAGAGTCCAGCACGCCCATCAAGGGCCCCCAAGAGG	1509
84	YHisProArgGlnAsnLeuHisIlySHisPheAspIleAsnGlnHisLeu	100
1510	GACATCTTAGACAGAAACCTACACCAAGCATTTTGCATCATGAATGAGCATTTG	1559
101	ProTrpMetIleValLeuPheLeuLeuLeuValLeuValIleValIleVal	117
1560	CCCTGATGATGATGCTCTTTCTCTGCTGGTGGCTTGGGATTTGTGGT	1609
117	LYsserIleArgLysSerSerArgThrLeuLysIlySGlyProArgGln	134
1610	GTCGACATATCCGGAAAGACTCGAGACATCTGTAAAAAGGGGCCCCGGCAGG	1659
134	sPProSerAlaIleValGlnIuLysAlaGlyLeuLysLysSerMetThrPro	150
1660	ATGCCAGTGTCCATTGTGGAAAGGCGAGGCTGAAGAATAATCCATGACTCCA	1709

151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
 |||
 1710 ACCGAGACCGGAGAAATGATCTACTACTGCAATGGCCATGTATCGA 1759
 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
 |||
 1760 TATCGTGAACCTGTGTAGACAGCCCAAGTGGAGCCAGTGGAAAGATATCT 1809
 184 YrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
 |||
 1810 ATCAGTTCTTTCATATGCCAGTGAAGAGGAGGTGCTGCTTCTCCAT 1859
 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
 |||
 1860 GGGTACACACAGCCAGCAGAGCGGGCTACGACGCTCTGCAGCAGCTGGAC 1909
 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
 |||
 1910 CATCCGGGGCCCGAGGCGCCTCGCCAGCTAATTAGCGCCCTGCGCC 1959
 234 InHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
 |||
 1960 AGCACCGGAGAAACGATGTTGTGAGAGAGATTGCTGGCTGATGAGAGAC 2009
 251 ThrThrGlnLeuGluThrAspLysLeuValAlaLeuProMetSerProSerPr 267
 |||
 2010 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCAAGCCC 2059
 267 cIeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
 |||
 2060 GCTTACCCCGAGCCCATCCCAAGCCCAAGCGGAAACTTGAGAAATTCCG 2109
 284 lAlaLeuLeuThrValGluProSerProGlnAsp 294
 |||
 2110 CTCTCTGACGCTGAGGCTTCCCAACAGGAT 2141

801 GAGTAGCATCCAGGAAGGACAGCTCCCTGAGAAACAAGCTCAGCAAGGG 850
 34 IYLYSLAAspValasnlystrleuProasnleuGluValValasnHs 50
 851 GGAAGGAGACGTGAAACAGACCTCCCAACCTTCAGGTAGTCAACAC 900
 51 GlnGlnGlyProHsHsArgHsIleleuLysleuLeuProSerMetC1 67
 901 CAGCAAGGCCCCACACAGACACATCTGAGAGCTGCTCCGTCATGGA 950
 67 uAlaThrGlyGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
 951 GCCCAGTGGGGGAGAGATCCAGCACGCCCATCAAGGGCCCCAAGAGG 1000
 84 IYH1PProArgGlnasnleuHsIlyHsHsPheAspIleasnGluHsleu 100
 1001 GACATCCCTAGACAGAACTACACAAGCATTGTCATCATGAGCATTTG 1050
 101 ProTrpMetIleValleuPheleuLeuValleuValIleValIle 117
 1051 CCTGTGATGATGTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
 117 IcySerIleArgLysSerSerArgThrleuLysLysGlyProArgGlnA 134
 1101 GTGCAGTATCCGGAAGAGCTCGAGAGCTCTGAAAAAGGGCCCCGCGAG 1150
 134 SPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
 1151 ATCCAGAGTGCATGTGGAAGGCGAGGCTGAGAAATCCATGACTGCA 1200
 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysasnGlyHsGlyIleAs 167
 1201 ACCCGAAGCCGGAGAAATGATGATCTACTACTCAATGGCATGATGCA 1250
 167 PlleuLysleuValAlaIleGluValGlySerGlnTrpLysAspIleT 184
 1251 TATCTGTAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATAT 1300
 184 YrGlnPheLysCysAsnAlaSerGluArgGluValAlaIlePheSerAsn 200
 1301 ATCAGTTTCTTTCGCAATGCACTGAGAGGAGGAGTGTCTTTCGCAT 1350
 201 GlyTyrThrAlaAspHsIleuArgAlaTyrAlaIleleuGlnHsTrpTh 217
 1351 GGGTACACAGCCGACGACGAGGCGGCTTACGACGCTGCGACACTGAC 1400
 217 rIleArgGlyProGluAlaSerleuAlaGlnleuIleSerAlaLeuArg 234
 1401 CATCCGGGGCCCCGAGGCGAGCTCGCCCAATTAATTAGCCCTGCGCC 1450
 234 IuHsArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
 1451 AGCACCGGAGAAACGATGTTGTGAGAAAGATTGCTGGGCTGATGGAAG 1500
 251 ThrThrGlnleuGluThrAspLysleuAlaLeuPrometSerProSerPr 267
 1501 ACCACCCAGCTGGAACGACAACTAGCTTCGCGATGAGGCCAGCCAGCC 1550
 267 oleuSerProSerProIleProSerProAsnAlaLysleuGluAsnSerA 284
 1551 GCTTAGCCCGAGCCCATCCCAAGCCCAAGCGGAACCTTGGAATTTCCG 1600
 284 IAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
 1601 CTCTCTGAGGCTGGAGCTTCCTCCGACAGACAAGAAAGGCTTCTTTC 1650
 301 ValAspGluSerGluProleuLeuArgCysAspSerThrSerSerGlySe 317
 1651 GTGGAATGAATCGAGAGCCCTTCTCCGCTGTGACTTACATCCAGCGGCTC 1700
 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysAspTr 334
 1701 CTCGCGGTGAGAGGAGACGTTCTTTATTCACAAAGAAAGAGGACA 1750

334 hrValleuArgGluValArgLeuAspProCysAspLeuGlnProIlePhe 350
 1751 CAGTGTGGGGAGGTAGCCCTCGAGACCCCTGTGACTTGAGGCTATTCCTT 1800
 351 AspAspMetLeuHsIlePheLeuAsnProGluGluLeuArgValIleGluC1 367
 1801 GATGACATGCTCCACTTCTTAATCTGAGAGAGCTGCGGCTGATGGAAGA 1850
 367 uIleProGlnAlaGluAspLysleuAspArgLeuPheGluIleIleGly 384
 1851 GATTCGCCAGGCTGAGAGACAACTAGACCGGCTATTGCAATTTATGGAG 1900
 384 aLysSerGlnGluAlaSerGlnThrleuLeuAspSerValTyrSerHs 400
 1901 TCAGAGGCGAGAGACGACGACGACCCCTCTGAGCTCTGTTATAGCCAT 1950
 401 LeuProAspLeuLeu 405
 1951 CTTCCTGACCTGCTG 1965

seq_name: gb_pat:E29452
 seq_documentation_block: 2186 bp DNA PAT 07-FEB-2001
 LOCUS E29452
 DEFINITION Tumor necrosis-associated receptor TR7.
 ACCESSION E29452
 VERSION E29452.1 GI:13025564
 KEYWORDS JP 1999018786-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 2186)
 AUTHORS Keith,C.D.M.M., Hurle,P.Y.K.R. and Tan.
 TITLE Tumor necrosis-associated receptor TR7
 JOURNAL Patent: JP 1999018786-A 1 26-JAN-1999;
 SMITHKLINE BEECHAM CORP
 OS Unidentified
 PN JP 1999018786-A/1
 PD 26-JAN-1999
 PE 02-APR-1998 JP 1998126559
 PR 02-APR-1997 US 60/041796, 28-OCT-1997 US 08/959382 P1
 KETTH C DEAN, MARK R HURLE, PETER YOUNG, K B TAN PC
 C12N15/09, A61K38/00, A61K38/00, A61K38/00, A61K45/00, PC
 A61K48/00,
 PC C07K14/705, C07K16/30, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC
 C12N5/10,
 PC C12P21/02, G01N33/53, G01N33/531, G01N33/566, G01N33/577,
 C12N15/00, PC A61K37/02,
 PC A61K37/02, A61K37/02, A61K37/02, C12N5/00, C12N5/00 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..2186
 FT Location/Qualifiers
 FT 1..2186
 /organism="Unidentified".
 /organism="unidentified"
 /db_xref="taxon:32644"

FEATURES
 source
 BASE COUNT 514 a 613 c 565 g 494 t
 ORIGIN

alignment_scores:
 Quality: 2091.00 Length: 405
 Ratio: 5.163 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x E29452 ..
 Align seg 1/1 to: E29452 from: 1 to: 2186


```

1 MetAsnSerThrgluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
757 ATGACATCAACAGATCCACACTCTTCTGCTGTGTAGACCAAGGACT 806
17 userSerIleGlnGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
807 GAGTACGATCCAGGAAGGAGACGTCCCTGACAAACACAAAGCTCAGCAAGG 856
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||
857 GGAAGAGAGACGTGACACACACACCTCCCAACCTTCAGTAGTCACACC 906
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG 67
|||||
907 CACAGAGGCCCCACACACACACATCTGAGCTGCTGCCGTCCATGGA 956
67 uAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
957 GGGCACTGGGGGGAGAACTCCAGCAGCCCATCAAGGCCCAAGAGG 1006
84 LHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1007 GACATCTTACAGAACCTACACAGCATTTTGCATCAATGAGCATTTG 1056
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleVala 117
|||||
1057 CCTGGATGATTTGCTTTCTCTGCTCTGCTGCTTGTGATTTGGGT 1106
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlna 134
|||||
1107 GTCAGATATCCGGAAGAGCTCGAGACTGTAAAGGGCCCGCAGG 1156
134 sProSerAlaIleValGlnLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1157 ATCCAGTGCATTGTGGAAAGCAGGAGCTGAAGAAATCCATGACATCCA 1206
151 ThrGlnAsnArgLysLysTrpIleTyrrTyrrCysAsnGlnHisGlyLeas 167
|||||
1207 ACCCAGAACCGGGAATGATGATCTACTACTGCAATGGCCATGGTATCGA 1256
167 PileLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
|||||
1257 TATCCGTGAACCTGTAGACAGCCCAAGTGGGAGACCACTGGAAGATATCT 1306
184 YrGlnPheLeuCysAsnAlaSerGlnArgGlnuValAlaIlePheSerAsn 200
|||||
1307 ATGAGTTTCTTGTCAATGCGAGTAGAGGAGGTGCTGCTTCTTCCAAT 1356
201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisIleTrpTh 217
|||||
1357 GGGTACACAGCCGACGACGAGCGGCTTACGCAAGCTTGCAGCAGCTGGAC 1406
217 rIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||
1407 CATCCGGGGCCCGAGGCGAGCTCGCCGCAATATGAGCCCTGCGCC 1456
234 LHisArgArgAsnAspValValGlnLysIleArgGlyLeuMetGlnuAsp 250
|||||
1457 AGCACCAGGAACGATGTTGTGAGAGAGATTCGTGGCTGATGGAAGAC 1506
251 ThrThrGlnLeuGlnThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1507 ACCACCAGCTGGAACTGACAAACTAGCTCTGCCGATGAGCCCAACCC 1556
267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGlnuAsnSera 284
|||||
1557 GCTTAGCCCAAGCCCATCCCAAGCCCAACGCAAACTGAGAAATTCCG 1606
284 lAlaLeuThrValGlnProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1607 CTCTCCCTGAGCGTGAGACTTTCCTCCACAGACAAAGAAAGGCTTCTTC 1656
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317

```

```

|||||
1657 GTGATGAGTCGAGACCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1706
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGlnLysLysAspT 334
|||||
1707 CTCGCGCTGTGAGCAAGAGGTTCCTTATTTACCAAGAAAGAAAGACCA 1756
334 hrValLeuArgGlnValAlaArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
1757 CAGTGTTCGGGAGGAGTACGCTTGAGACCCCTGTGACTTGCACCTATCTTT 1806
351 AspAspMetLeuHisPheLeuAsnProGlnGlnLeuArgValIleGlnG 367
|||||
1807 GATGACATGCTCCACTTTCTTAATCTGAGAGAGCTCGCGGCTGATTGAAGA 1856
367 uIleProGlnAlaGlnAspLysLeuAspArgLeuPheGlnIleIleGly 384
|||||
1857 GATTCCCGAGGCTGAGGACAAACTAGACCGGCTATTGCAAAATTAATTGGAG 1906
384 aLysSerGlnGlnAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
|||||
1907 TCAAGACCCAGGAAGCCAGCCAGACCTCTGAGACTGTATTATAGCAT 1956
401 LeuProAspLeuLeu 405
|||||
1957 CTTCCTGACCTGCTG 1971
seq_name: gb_pat:AX008646
seq_documentation_block:
LOCUS AX008646 2666 bp DNA PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO9966039.
ACCESSION AX008646
VERSION AX008646.1 GI:9996171
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2666)
AUTHORS Fields,W.E., Gomes,B.C. and Lu,J.J.
TITLE Trif-like protein with death domain
JOURNAL Patent: WO 9966039-A 1 23-DEC-1999;
ZENCA LTD (GB); FIELDS WILLIAM EDWARD (US); GOMES BRUCE CHARLES
(US); LU JIAN JASON (US)
FEATURES
location/Qualifiers
source
BASE COUNT 691 a 609 c 591 g 775 t
ORIGIN
alignment_scores:
Quality: 2091.00 Length: 405
Ratio: 5.163 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-733-956-2 x AX008646
Align seg 1/1 to: AX008646 from: 1 to: 2666
1 MetAsnSerThrgluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
190 ATGACATCAACAGATCCACACTCTTCTGCTGTGTAGACCAAGGACT 239
17 userSerIleGlnGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
240 GAGTACGATCCAGGAAGGAGACGTCCCTGACAAACACAAAGCTCAGCAAGG 289
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||

```

230 GGAAGAGACGTGAACAGACACCTCCCAACCTTCAGGTAGTCAACAC 339
 51 GINGINGLYPRHISHSISARHISITILEULYSLEULEUPROSERMETG 67
 340 CAGCAAGGCCCCACCACAGACACATCTGAAAGCTCTCCCTCCATGGA 389
 67 uAlaThrGlyGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
 390 GCCCAGCTGGGGCGAGAAAGTCCAGCACGCCCATCAAGGGCCCCAGAG 439
 84 LYSIPROARGINASNLEUHSISHSISPHASPLIANSNGLNHSLEU 100
 440 GACATCTAGACAGACCTACACAAAGCATTTTGACATCATATGACATTGG 489
 101 PROTPMETILEVALLEUPHEULEULEUVALLEUVALILEVALYA 117
 490 CCCGTGATGATCTGTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 117 LCYSERILEARGLYSERSERARGTHRLEULYSGLYPROARGLNA 134
 540 GTGCAGTATCCGGAAGAGCTCGAGACCTGAAAAAGGGCCCCGGCAG 589
 134 SPPTOSERLALILEVALGILYSLAIGLYLEULYSSETMETHPRO 150
 590 ATCCAGTGTCCATTTGTGAAAGGAGGCTGAAGAAATCCATGACTCCA 639
 151 THRGINASARGILULSTPRILETYTYTYCYASNGLYHISGLYLLEAS 167
 640 ACCCAAGACCGGGAGAAATGATGATCTACTCTCATGSCCATGTATGCA 689
 167 PILEULEULYSLEUVALLEALAGLVALIGLYSERGINTPLYSAPLLET 184
 690 TATCTGTGAAGCTTGAGAGCCAGTGGGAAAGCCAGCTGAAAGATATCT 739
 184 YGILPHLEUCYSASNALISERGLUARGLVALALALAPHERASRN 200
 740 ATCAGTTCTTTTGCATGTCCAGTGAAGAGGAGTTCCTCTTCTTCAT 789
 201 GLYTYRTHALIASPHISGLUARGLALATYRALALALEUGLNHSITRPH 217
 790 GGGTACACAGCCGACACAGAGGGGCTACGACGCTCTCAGACACGAG 839
 217 TLEARGGLYPROGLUALASERLEUALAGLLEULEULEULSERLALALEUARG 234
 840 CATCCGGGGCCCCGAGGCCAGCTCCGCAATATACGCCCTCCGCC 889
 234 LNHISARGASNASPVALLAIGLULYSILEARGLYLEUWETGUASP 250
 890 ACCACCGGAGAAACGATGTTGTGAGAGATTCGTGGCTGATGAGAGAC 939
 251 THRTHGLNLEUGLUTHRASPLYSEUALALEUPROMETSERPROSERPR 267
 940 ACCACCCAGCTGGAACGTACAACAGCTCTCCGATGAGCCCCAGCCC 989
 267 OLEUSERPROSERPROLIEPROSERPROASNALALYSLEUGLUNASERA 284
 990 GCTTAGCCCGAGACCCCAACCCAGCCCAACCGGAAACTTGAGAAATTCG 1039
 284 LALEULEUTHRALIGUPROSERPROGLNASPLYSNLSGLYPHEPHE 300
 1040 CTCCTCTGACGGTGGAGCCTTCCCCACAGACACAGACAGAGGGCTTCTTC 1089
 301 VALASPGULSERGLUPROLEULNARGCYASPSERTHSETSERGLYSE 317
 1090 GTGGATGATGAGAGCCCTCTCCGCTGTGACTACATCCAGCGGCGTC 1139
 317 ISEIRLALAEUSERARGASNGLYSERPHEILETHRLYSGLULYSASPT 334
 1140 CTCGCCGCTGAGACAGAGAGGTTCTTTATACCAAGAAAGAAAGAGACA 1189
 334 hVALLEUARGLINALARGLEUASPPOCYASPLEUINPROLIEPHE 350
 1190 CAGTGTGGCGAGGTACGCTGAGACCCCTGTGACTTGTGACGCTATCTTT 1239

351 ASPASPMELLEUNHSIPHELEUNASPROGLUGLULEUARGVALILEGLUG 367
 1240 GATGACATCTCTCACTTCTTAATCTGAGAGAGCTCGGGGTGATTGAGA 1289
 367 uLLEPROGLINALAGLUSPLYSLEUNASPARGLUEPHEGLUILEIGLY 384
 1290 GATTTCCCGAGCTGAGACAACTAGACCGGCTATTGCAAAATATATGGAG 1339
 384 ALLYSERGLINGLUALISERGINTHLEULEUNASPSERVALTYRSEHIS 400
 1340 TCAAGAGCCAGAGAGCCAGACACCTCTCTGAGACTCTGTTATAGCCAT 1389
 401 LEUPROASPLEULEU 405
 1390 CTTCCTGACCTGCTG 1404

seq_name: gb_pr:AK001504

seq_documentation_block:

LOCUS AK001504 2636 bp mRNA PRI 22-FEB-2000

DEFINITION Homo sapiens cDNA FLJ10642 fis, clone NT2RP2005752, highly similar

to Homo sapiens TNFR-related death receptor-6 mRNA.

ACCESSION AK001504.1 GI:7022799

VERSION AK001504.1

KEYWORDS oligo capping; fis (full insert sequence).

Source Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,

clone.lib:NT2RP2 clone:NT2RP2005752.

ORANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (sites)

Isoqai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,

Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,

Tanai,H., Kinata,M., Matanabe,M., Hirooka,S., Ishii,S., Kawai,Y.,

Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K.,

Masuno,Y. and Kanemori,K.

NEO human cDNA sequencing project

Unpublished (2000)

Isoqai,T. and Otsuki,T.

Direct Submission

Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao

Isoqai, Helix Research Institute, Genomics Laboratory: 1532-3 Yana,

Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,

Tel:81-438-52-3951, Fax:81-438-52-3952)

NEO human cDNA sequencing project supported by Ministry of

International Trade and Industry of Japan; cDNA full insert

sequencing: Research Association for Biotechnology; cDNA library

construction, 5'- & 3'-end one pass sequencing and clone selection:

Helix Research Institute (supported by Japan Key Technology Center

etc.) and Department of Virology, Institute of Medical Science,

University of Tokyo.

FEATURES

1. 2636

Location/Qualifiers

1. 2636

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="NT2"

/cell_type="teratocarcinoma"

/clone="NT2RP2005752"

/clone_lib="NT2RP2"

/note="Cloning vector: pME18SFL3; mRNA from NT2 neuronal

precursor cells after 2-weeks retinoic acid (RA)

induction."

BASE COUNT 569 a 762 c 722 g 583 t

ORIGIN

alignment_scores:

Quality: 2087.00 Length: 405

Ratio: 5.153 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.753

alignment_block:
US-09-733-956-2 x AK001504

Align seg 1/1 to: AK001504 from: 1 to: 2636

```

1 MetAsnSerThrgluSerAsnSerSerAlaSerValArgProLysValLe 17
|||||
1145 ATGAACCTCAACAGAACTCAACTCTTCTGCTCTGTTAGACCAAGGATCT 1194
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
1195 GAGTAGCATCCAGAGAGGAGGACAGTCCCTGACACACAAAGCTCAGCAAGG 1244
34 LylYsGluAspValAlaSnLysThrLeuProAsnLysGlnValAlaSnHis 50
|||||
1245 GGAGAGAAAGAGCTGACACAGACCTCCCAACCTTCCAGTAGTCACACAC 1294
51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||
1295 CAGCAAGGCCCCCACCACAGACACATCCTGAAAGCTGCGCGTCATGGA 1344
67 uAlaThrglyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
1345 GGCCACGCGGGGCGAGAGTCCACAGCCCATCAAGGCCCCCAAGAGG 1394
84 LysIleProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1395 GACATCTCTAGACAGAACCTACACAAAGCATTTTGACATCAATGACATTTG 1444
101 ProTrpMetIleValLeuPheLeuLeuValLeuValIleValLeu 117
|||||
1445 CCCGTGATGATGTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
117 LysSerIleArgLysSerSerArgThrLeuLysGlyProArgGln 134
|||||
1495 GTGAGATATCCGGAAAGCTCGAGACTGTGAAAGGGGCGCGCAGG 1544
134 sPrProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
1545 ATCCCACTGCCATTGTGGAAGAGCAGGCGCTGAAGAAATCCATGACTCCA 1594
151 ThrGlnAsnArgLysLysTrpIleTyrTyrCysAsnGlnHisGlyLeu 167
|||||
1595 ACCCAGAAACCGGAGAAATGATCTACTACTGCAATGCCCATGTGATCGA 1644
167 PileLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
|||||
1645 TATCTTAAGCTGTGACAGCCCAAGTGGGAAGCAGTGAAGATATATCT 1694
184 YrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
|||||
1695 ATCAGTTCTTTTGCATGCCAGTGAGAGAGGTTGCTGCTTTCTCCAT 1744
201 G1YrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
|||||
1745 GCGTACACAGCCGACACAGAGCGGCGCTACACACTGTCAGCACTGAC 1794
217 rIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||
1795 CATCCGGGGCCCCGAGGCCACGCTGCCCACTAATTAGCCGCTGCGCC 1844
234 LnhIleAspArgAsnAspValAlaGluLysIleArgGlyLeuMetGluAsp 250
|||||
1845 AGCACCGGAGAAACGATCTTCTGAGAAAGATTCTGGGCTATATGCAAGC 1894
251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1895 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCCAGGCC 1944
267 oLeuSerProSerProIleProSerProAsnAlaLysLeuLysAsnSer 284
|||||
1945 GCTTAGCCCGGCCCCCATCCACGCCCAACGCAAAACTTGAGAAATTCGG 1994

```

```

284 lAlaLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
|||||
1995 CTCTCTGAGCGGTGGAGGCTTCCCCACAGACAGAAACAGAGGCTTCTTC 2044
301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
|||||
2045 GTGATGAGTGGAGAGCCCTTCTCCGCTGTGACTCTACATCCAGGCGCTC 2094
317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysAspT 334
|||||
2095 CTCCGCGCTGAGCAGGAGCGGTCTCTTATTACCAAAAGAAAGAGCA 2144
334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
|||||
2145 CAGTGTTCGGCAGGTACGCTCGACCCCTGTGACTTCAGAGCTTATCTTT 2194
351 AspAspMetLeuHisPheLeuAsnProGluLysLeuArgValIleGluG1 367
|||||
2195 GATGACATGCTTCACACTTCTAATCTGAGAGGTGCGGGTGATGAGGA 2244
367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleGly 384
|||||
2245 GATTCCCGAGGCTGAGGACAACTAGACCGGCTATTGCAAAATTATTGGAG 2294
384 aLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
|||||
2295 TCMAGAGCCAGGAAGCCAGCCAGACTTCTGTGACTCTGTTATTAGCAT 2344
401 LeuProAspLeuLeu 405
|||||
2345 CTCTCTGACTGCTG 2359

```

seq_name: gb_pat:AR134329

seq_documentation_block:

LOCUS AR134329 1719 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 5 from patent US 6194151.

ACCESSION AR134329

VERSION AR134329.1 GI:14123234

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

alignment_scores:

Quality: 1914.00 Length: 406

Percent Similarity: 96.552 Ratio: 4.883

Percent Identity: 92.857 Gaps: 2

alignment_block:

US-09-733-956-2 x AR134329

Align seg 1/1 to: AR134329 from: 1 to: 1719

1 MetAsnSerThrgluSerAsnSerSerAlaSerValArgProLysValLe 17

|||||

505 ATGAACCTCAACAGATTCACACTTACTGCTCTGTTAGAACTAAGGTACC 554

17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34

|||||

555 AAGTGCATTCGAGAGAGGACAGTCTGACAAATAGAGCTCAACACACTG 604

34 LylYsGluAspValAlaSnLysThrLeuProAsnLysGlnValAlaSnHis 50

605	GSAGAGGACGACACTATATGAGACGACCGCCAAACCCAGCCAGATTACCCAC	654
51	gInglIngLyProOnHsHsArghIsIleuLysLeuLeuPro...SerMe	66
655	CAGCAAGGCCCCCACCACAGACACACTTCTGAAGGCTCGCCATGGTCCAT	704
66	tGluLhArhrgLygLyLysSerSerThrProLLeysGlyProLysA	83
705	GGAGGCCACG...GGTGGAAGTCCAGCAGACGCCATCAAGGCCCCCACA	751
83	rgLhIsProArgrgLnAsnLeuHsHsLysHsPhAspDLeaNgLhIs	99
752	GGGGTCACCCCAAGAGAACGCTGCACAGCATTTGCAATCATCAGCAGCAC	801
100	LeuProTrpMetIleValLeuPheLeuLeuValLeuValIleVal	116
802	TTGCCTTGAGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	851
116	lValLysSerIleArgLysSerSerArgThrLeuLysLysGlyProArg	133
852	GGTGCGAGTATCCCAAGAGCTCCAGGACCTCCAAAAGGAGGCCCGCG	901
133	lAsProSerAlaIleValGlyLysAlaGlyLeuLysLysSerMetThr	149
902	AGGATCCACGCCCCATAGTGGAAAAGCGGGGCTGAAGAATGCTGACT	951
150	ProThrGlnAsnArgGlyLysTrpIleTyrTyrCysAsnGlyHsGlyTl	166
952	CCCAACCCAGAACCGGGAGAAATGCATCTACTAGCCCAAGGCCATGCTAT	1001
166	eAspIleuLysLeuValAlaIleGlnValGlySerGlnTrpLysAspT	183
1002	TGACATCTTGAAGCTTGTAGACGCCAGCTGGGAAGCCAGTGGAAAGCA	1051
183	lEtyrGlnPheLeuLysAsnAlaSerGlyValrGlnValAlaAlaPheSer	199
1052	TCATACAGTTCTTTGGCAACGCCACTGAGAGAGAGTGGCGCTTCTCC	1101
200	AsnGlyTyrThrAlaAspHsGlyLysrGlyAlaTyrAlaAlaLeuGlnHsTr	216
1102	ATGGAATACACGCGAGATCATGAAACGGGCGCTACGCGCTCGACACATG	1151
216	pThrIleArgGlyProGlnAlaSerLeuAlaGlnLeuIleSerAlaLeuA	233
1152	GACCAATCCGTGGCCCTGAGGCCAGCCTTGGCCAGCTCATTAAGCGCTTGC	1201
233	rGlnHsIArgArgAsnAspValValGlyLysIleArgGlyLeuMetGlu	249
1202	GCCACACCCGAGGCATATGTTGTGGGAAGAAATTCTGGGGCTGATGGAA	1251
250	AspThrThrGlnLeuGlyThrAspLysLeuAlaLeuProMetSerProSe	266
1252	GACACACAGCAATGGAAACAGCAAACTGGGCTGCCCATGAGCCCCAG	1301
266	rProLeuSerProSerProIleProSerProAsnAlaLysLeuGlnAsnS	283
1302	TCGCGTGAAGCCGAGACCCCATCCCGAGTCTTAACGTCAAACTTGGAGATT	1351
283	eAlaIleuLeuThrAlaGlyProSerProGlnAspLysAsnLysGlyPhe	299
1352	CCACGCTCTCGACATGGAGCCCTCACCGCTGGAGAAACAAAGAGCTTC	1401
300	PheValAspGlnSerGlyProLeuLeuLysCysAspSerThrSerSerGly	316
1402	TTTCGTGGACAGATCAGAGCCCTTTGGCTTGGCGACTCACATCCAGTGG	1451
316	YSerSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLysLysA	333
1452	CTCTTCAGACGTCAGACAGAAAGGCTCTTATATTACAAAGAAAGAAAGG	1501
333	spThrValLeuArgGlnValAlaGleuAspProCysAspLeuGlnProIle	349

```

1502 ACACAGTGTTCGGCAGAGTCCGGCTGGACCCCGTCGATCTTGGACGCCATTC 1551
350 pheaspaspmetleuHispheleuasnProGluIleuAtrValIleGI 366
|||||.....
1552 TTGTGTGATCATGTGCTGATATCTGAAACCCCGAGACGACTCCGGGTATGA 1601
366 uGluIleProGlnIlaGluaspLysleuaspArgleuPheGluIleIleG 383
|||||.....
1602 ACAGATTCCTCCAGGCTGAGGACAACTGACGCCCTCTTCGAGATCATTTG 1651
383 lyaLlysSerGlnGluAlaSerGlnThrleuLeuaspSerValTyrSer 399
|||||.....
1652 GGGTCAGACGCCAAGAGCCAGACGCCCTCTTGGACTCTGTACAGT 1701
1702 CATCTTCCTGACTATTTG 1719
seq_name: gb_pat:AR134327
seq_documentation_block:
LOCUS AR134327 3331 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6194151.
ACCESSION AR134327
VERSION AR134327.1 GI:14123232
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3331)
TITLE Busfield,S.J.
FEATURES Molecules of the TNF receptor superfamily and uses therefor
source Patent: US 6194151-A 1 27-FEB-2001;
Location/Qualifiers
1.3331
source "organism="unknown"
BASE COUNT 825 a 891 c 739 g 876 t
ORIGIN
alignment_scores:
Quality: 1914.00 Length: 406
Ratio: 4.883 Gaps: 2
Percent Similarity: 96.552 Percent Identity: 92.857
alignment_block:
US-09-733-956-2 x AR134327 ..
Align seg 1/1 to: AR134327 from: 1 to: 3331
1..MetasSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||.....
848 ATGACCTCAACAGATTCACACTGACCTCTGTAGAACTAAGTACC 897
17 userSerIleGlnGluGlyThrValProaspasnThrSerSerAlaArg 34
|||||.....
898 AAGTGGCATTCGAGAGGAGGACAGTGGCTGCAATTAAGACTCAACACAGT 947
34 lylGluaspValAlasnlserThrleuProaspasnleuIValValasnhIs 50
|||||.....
948 GGAAGGAAGGCACTAATTAAGACCTGGCAAAACCCACCAAGTTACCCAC 997
51 GlnGlnGlyProHnHisArgHisIleLeuLysleuLeuPro...SerMe 66
|||||.....
998 CAGCAAGCCCCCACACACAGACATTTCTAAGCTGCTGCATGCTCAT 1047
66 lGluAlaThrIleGlyGluLysSerSerThrProIleLysGlyProLysA 83
|||||.....
1048 GAGAGCCACAG...GGTGAAGATGCCACACAGCATCAAGGCCCCACAGA 1094
83 rGlyHisProArgGlnaspLeuHisLysHisPheaspLysasnGlnHis 99
|||||.....
1095 GGGGTACCCCGACAGAGAGGCTTCACAAAGCATTTGACATCAACAGAGCAG 1144

```



```

34  lYlYsGIuAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
      |||||  ::::::::::|
851  GGAAGAGAGCACTAATAGACCCCTGCCAAACCCACACAGTTACCCAC 900
51  GInGInGlyProHisHisArgHisIleLeuLysLeuLeuPro...Serne 66
      |||||  ::::::::::|
901  CAGACAGCCCCACCAAGACACATTTGAAAGCTCTGCTGCTGCTGCTGCT 950
66  tGluAlaThrGlyGlyGlyLysSerSerThrProIleLysGlyProLysA 83
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
951  GAGAGGCCAAG...GCTGAGAGTCCAGCACAGCCATCAGAGGCCCAAGA 997
83  rGgIYHISProArGgIuAsnLeuHisLysHisPheAspIleAsnGluHis 99
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
998  GGGGTCCACCCAGACAGAACGGCTCACACACATTTGCACATCAAGAGCAC 1047
100  LeuProThrMetIleValLeuPheLeuLeuValLeuValIleVal 116
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1048  TTGCTTGGATGATGCTGCTCTTCTCTGCTGCTGCTGCTGCTGCTGAT 1097
116  lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArg 133
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1098  GGTGTGCAATGCCGAAGAGCTCCAGACTCTCAAAAGGGGGCCCCGGC 1147
133  lAsPProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1148  AGGATCCAGCCCATGTGGAAGAGCGGGGCTGAAGAAAGTCCCTGACT 1197
150  ProThrGlnAsnArgGlyLysThrIleTyrTyrCysAsnGlyHisGlyIle 166
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1198  CCCACCCAGAACCGGAGAAATGATCTACTACCGACAGCGCCATGCTAT 1247
166  eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTyrPlyAspI 183
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1248  TGACATCTTGAAGCTGTAGACAGCCAGGTGGAGACCAAGTGAAGACA 1297
183  letyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1298  TCTATCAGTTCTTGGCAAGCCAGCGAGAGGGAGGTGGCGCTCTGCC 1347
200  AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1348  AATGATACACTGCAGATCATGACGGGCTTACGGGCTCTGCAGACACTG 1397
216  pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1398  GACCATCCCTGGCCCTGAGGCCAGCTTGCCTCCAGCTCATTTAGCGCTTG 1447
233  rGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1448  GCGAGACCGGAGCAATGATGTGTGGAAGAAGATTGTTGGCTGATGGAA 1497
250  AspThrThrGlnLeuGlnThrAspLysLeuAlaLeuProMetSerProSe 266
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1498  GACACACAGCACTTGGAAACAGCAAACTGGCTCTCCCATATGAGCCCA 1547
266  rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsn 283
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1548  TCCGCTTACCCGAGGCCATGCTCCAGTCTTAACGAAACTTGAAGATT 1597
283  eAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1598  CCAGCTCTCTGACAGTGGAGCCCTCACCGCTGACAAAGAACAGTCTTC 1647
300  PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerG 316
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1648  TTCGTGGAGAGATCAAGACCCCTTCTGCATGCACTCCACATCCAGTGG 1697
316  ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGlyLysA 333
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1698  CTTCTCAGACTGAGCAAGAGCGCTCTTTATTACCAAGAAAGAAAGAG 1747

```

```

333  spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1748  ACACAGTGTGGCGGAGGTCGCCGCTGAGCCCTGTGACTTGCAGCCCATC 1797
350  PheAspAspMetLeuHisPheLeuAsnProGlnGluLeuArgValIleG 366
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1798  TTTGATGACATGCTGATATCTGTAACCCCGAGAGACTGCGGTGATGA 1847
366  uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIle 383
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1848  AGAGATTTCCCGACGCTGAGAGCAAACTGAGCCGCTCTTGCAATCATTC 1897
383  lYValLysSerGlnGluAlaSerGlnThrLeuLysSerValTyrSer 399
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1898  GGGTCAAGAGCCAAAGAGCCAGACACCTTGTGACTCTGTACAGT 1947
400  HisLeuProAspLeuLeu 405
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1948  CATCTTCTGACCTATTG 1965
seq_name: gb_AF322069

```

```

seq_documentation_block: 1968 bp mRNA ROD 06-DEC-2000
LOCUS AF322069
DEFINITION Mus musculus DR6 mRNA, complete cds.
ACCESSION AF322069
VERSION AF322069.1 GI:11559849
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.

```

```

REFERENCE
AUTHORS Isogai,D., Ichino,M., Yoshinari,M., Yamaura,A., Kurokawa,F. and Minami,M.
TITLE Mouse DR6: mouse homolog of human TNFR-related death receptor-6 (DR6)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1968)
AUTHORS Isogai,D., Ichino,M., Yoshinari,M., Yamaura,A., Kurokawa,F. and Minami,M.

```

```

TITLE JOURNAL
SUBMITTED (17-NOV-2000) Immunology, Yokohama City University School of Medicine, 3-9 Fukuura, Kanazawa-Ku, Yokohama, Kanagawa 236-0004, Japan
FEATURES
source
Location/Qualifiers
1..1968
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/tissue_type="kidney"
1..1968
/codon_start=1
/product="DR6"
/protein_id="AAG38115.1"
/db_xref="GI:11559850"

```

```

translation="MGTASSTITALASCSRTAGVGATWAGSLLLGLPSTTAQPE
OKTILSLPGYRHHVDRTGOVLTCDKCPAGTYVSHCNMSLRVSSCPAGFTPHENG
IERCHDCSOPCPWMIERLPCALTLDRCLCPPEMYSONGCAHTYCPVAGVARKG
TENEDVRCOKCARCTFSDVPSSVMKCAHTPCLONLNVYKPGKEDNVCGLFPS
STNPSGTYVFSHMEHSHDVPSTYEPDGMSSNPSASTYRTVPVSGIEGTPG
DNSTSGKESTNNTLPNPQVTHQQAAPHRHILKLPSSMEATGEKSTRAIKAPKG
HPRONAKHEDINELDLPLMTLVLLVLVLSGIRKSSFTLAKGPRDPSAIVKA
GLKSLPTONREKWIYRNGHIDILKLVAAAGSQWKDIYPLCNASEREVAFSN
GYTADHERAYALQHWITRCPEASLAOLISALRQRRNDVYEKIRGLMEPTOLETK
LALPMSPLSPSPSPSPNYKLENSLTLYVPPSLDKNKCPEVDESPDLRCOSTSG
SSALSNGSFTTRKKOTVLRQVLDPCDDLPITDDMLHIINPELVIERIIPAEK
LDRLEFETIGVKSQDSQTLDDSVYSHLPDL"

```

```

BASE COUNT 475 a 594 c 514 g 385 t
ORIGIN
alignment_scores:

```



```
/product="death receptor 6"
/protein_id="AAK29666.2"
/db_xref="GI:14043015"
/translation="MGAIVLAVALPDLPLFLGTADAPKLTSEONAVSLPAGKYHLNDR
ATNDELCDKCPAGTGVSKHVTCTLRGSCPPGCTTKHENGIERCHPOKPCBELPM
IEKTHCALDRECTGCTSGFQINDCVPTVGVGVKKGKGTETDVRKRLRET
FSDVPSSVMKCKRTYDCFGKNMNVKPGTESDVAQSPSLPNTSLTSSPAQMDGET
YEAPPTAVLPKGLNSVFDLSSSPAPVNGTAEPVNDTSANGTAVGASLSSAG
TAGQOSYRHRKHTSQAMKQPADEMAGKESLIPYRPRKGPVHGFIDNEHLPM
IVLELLVLVIVVCSYKRSRLKPRDDPSLIVEKAIKMKSTTPTONREKWTYC
NGHEDILCTGPAVAVGSMQDIYOLFNASREVAAASNGYAADHERAYALAHWTIR
GPEASLIOLISALROHROANDVEKIRGLMETPTPVOMQPOMOQDCSNDQCKLEMDL
ALPYSPEPLSPVTPSPKPPDAALVITVEPSESEKCFVDESEPLTCDSTSSSSSL
SRGSEFTTEKKKDVLRKVRDPDQPIEDMDHLLINPELHVIETPOAEDKLDRL
FEIAGVSQESQSLTLDVYSHLPDL"
```

BASE COUNT 661 a 853 c 762 g 598 t 5 others
ORIGIN

alignment_scores:
Quality: 1564.00 Length: 426
Ratio: 4.34 Gaps: 5
Percent Similarity: 84.507 Percent Identity: 74.883

alignment_block:
US-09-733-956-2 x AF349908 ..

Align seg 1/1 to: AF349908 From: 1 to: 2879

```
1 MetAspSerThrGluSerAsnSerAlaSerValArgProLysValLeu 17
:::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
691 CTGAACTCTCGCTTTTATGATCTCTCTCCAGCGCGCGGTGTGTC 740
17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
741 CATGGCACAGCAGCAGCAACAGTGCATACATGACACTCGGCCAACG 790
34 yLysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
791 GCACAGTTGTGCTCTCTGGAGCGCTCCAGCGCGCGCACCGCTGCCAG 840
51 GlnGlnGlyProHisIleArgHisIleLeuLysLeuLeu.....Pr 64
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
841 GCACAGAGCTACCGGCACAGCAGCACCAGCATGAGGCGACAGCACC 890
64 oSerMetGluAlaThrGlyGlyLysSerSerThrProIleLysGlyP 81
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
891 GCGCGCAGAGATGGCCGGGGGTGAGAAGTCCAGCATCCCTACAGCGCG 940
81 rGlySerGlyHisProArgGlnAsnLeuHisLysHisSphaSpIleAsn 97
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
941 CCCGGAGAGGCGCGGCC.....AACGTCCACCAGCATTCGACATCAAT 984
98 GluHisLeuProThrMetIleValLeuPheLeuLeuValLeuValAla 114
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
985 GAACACTTCCCTCGATGATCGCTCTTCTGCTGCTGCTGCTGCTGCTG 1034
114 lIleValValCysSerIleArgLysSerSerArgThrLeuLysLysGlyP 131
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1035 CATCGTGTCTGAGCGGTGAGAGAGCTCAGCGACTCTGAAGAAAGGAC 1084
131 rArgGlnAspProSerAlaIleValGlnLysAlaGlyLeuLysLysSer 147
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1085 CCAGCGCAATATCCAGTGCATCGTGAAGAAAGCTATTATGAAGAAATCC 1134
148 MetThrProThrGlnAsnArgGluLysThrPileTyrTyrCysAsnLysH 164
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1135 ACCACCCCCACAGACAGGAGGAAATGATCTACTACGCAACGGGCA 1184
164 sGlyIleAspIleLeuLysLeuValAlaIleGlnValGlySerGlnTrpL 181
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1185 TGGCATCGACATCTGTAAGCGCTGTGACGCCCGAGGTGGGAGCATGTGA 1234
```

```
181 yAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAla 197
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1235 AGGACATCTACACAGTCTCTGTCATTAAGTCCAGAGCGGAGGTGGCGCC 1284
198 pSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuG 214
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1285 TTCTCCAAAGGCTACGCGGCGCCAGCACAGAGCGGCGCTACGCGGTGCA 1334
214 nHsTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1335 GCACGTGACCATTCGCGGCGCTGAGGCGCAGCCTGGACAGCTCATCAGCG 1384
231 lAlaLeuArgGlnHisArgArgAsnAspValAlaGlnLysLleArgGlyLeu 247
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1385 CCTCCGCGCAGACCGCCCGCAACGAGCTGTGTGAAAGATCCGCGCGCTG 1434
248 MetGluAspThrThr..... 252
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1435 ATGGAGGACACACCGCGCGGTGCATGATGACGCTCACTGGCAAAACAGCA 1484
253 .....GlnLeuGlnThrAspLysLeuAlaLeuProm 263
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1485 CTGCTCAATGATGACGGAACCTGGAAGGTGACAAACGCGACTGCTTG 1534
263 eSerProSerProLeuSerProSerProIleProSerProAsnAlaLys 279
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1535 TGAGCCCGCAGCCCGCTCAGCCCGGTGCCACCCCGCAGCCCGC....AAA 1578
280 LeuGluAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAs 296
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1579 CCACCCGATGCGCGCTCTCTACCGTGGAGCCCTCTCTTCAGAGAGAG... 1626
296 nLysGlyPhePheValAspLysSerGluProLeuLeuArgCysAspSerT 313
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1627 .AAATGCTTCTTTGTGAGAGAGTGGAGCCCTGCTGCGGTGACTGCA 1675
313 hSerSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLys 329
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1676 CTTCCAGCGCGCTCTCTGCGTCCAGCGGAGCGCGCTCTTATTACCAAA 1725
330 GluLysLysAspThrValLeuArgGlnValArgLeuAspProCysAspLe 346
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1726 GAAAGAGAGACACAGTCTTGGCCGAGGTGCGCTTGACCTTGACCT 1775
346 uGlnProIlePheAspAspMetLeuHisPheLeuAsnProGluGluLeu 363
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1776 GCAGCCATCTTCGATGACATGCTGCACATCTCAACACCCGAGAGACTCC 1825
363 rGlyAlaIleGluGluIleProGlnAlaGluAspLysLeuAspArgLeuPhe 379
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1826 ACGTGATCGAGAGATCCCGCAGCGGAGGACAAAGCTTGACAGGCTGTT 1875
380 GluIleIleGlyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSe 396
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1876 GAGATGCGCGGAGTCAAGAGCCAGAGAGCCAGCAACCTCTCTGACTTC 1925
396 rValTyrSerHisLeuProAspLeuLeu 405
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1926 TGTCTATAGCAACCTCTGACTTACTG 1953
```

seq_name: gb_pat:ARI34340

seq_documentation_block:

LOCUS ARI34340 1815 bp DNA 16-MAY-2001

DEFINITION Sequence 24 from patent US 61,941,51.

ACCESSION ARI34340

VERSION ARI34340.1 GI:14123245

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1815)

AUTHORS Busfield, S.J.

TITLE Molecules of the TNF receptor superfamily and uses therefor
JOURNAL Patent: US 6194151-A 24 27-FEB-2001;
FEATURES Location/Qualifiers
SOURCE 1.1815
/organism="unknown"
BASE COUNT 448 a 529 c 462 g 376 t
ORIGIN

alignment_scores:
Quality: 1527.00 Length: 294
Ratio: 5.194 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AR134340 ..

Align seg 1/1 to: AR134340 from: 1 to: 1815

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
751 ATGACACACAGACAGATCCACTTCTGCTGCTTGTAGACCAAGTACT 800
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
801 GAGTAGCATCCAGAGAGGACAGTCCCTGACACACACAGCTCAGCAAGG 850
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||
851 GGAAGGAGAGCGTGACACAGACCTCCCAACCTTCAGTAGTCAACACC 900
51 GlnGlnGlyProHisArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||
901 CAGCAGGCCCCCAGCAGACACATCTGAGCTGCTGCCCTCATGGA 950
67 uAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
951 GGCACACTGGGGGAGAAAGCTCAGACAGCCCATCAAGGGCCCAAGAGG 1000
84 LysProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1001 GACATCTTACACAGAACCTACACAGCATTTTGCATCAATGACATTGG 1050
101 ProArgMetIleValLeuPheLeuLeuValLeuValIleValIle 117
|||||
1051 CCTGGATGATGTGCTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
117 LysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 134
|||||
1101 GTGCAGTATCCGGAAAGCTCAGACACTGTGMAAAGGGCCCGGCAGG 1150
134 sProSerAlaIleValGlnLysLysLysLysSerMetThrPro 150
|||||
1151 ATCCCACTGCGCATTTGGGAAAGCGAGGCTGGAATAATCATCATCA 1200
151 ThrGlnAsnArgGlnLysGlyPheLeuTyrcysAsnGlyHisGlyIleAs 167
|||||
1201 ACCCAGAACCGGGGAAATGATCTACTACTGCAATGGCCATGTGATCGA 1250
167 PheLeuLysLeuValAlaIleGlnValGlySerGlnTrpLysAspIle 184
|||||
1251 TATCTGACAGCTTGTAGCAGCCCTGTGGGAAGCAGCATGGAACATAT 1300
184 TyrGlnPheLeuGlyAsnAlaSerGluArgGlnValAlaIlePheSerAsn 200
|||||
1301 ATGAGTTTCTTTGCAATGCGAGTGAAGGAGGTGCTGCTTTTCCAAAT 1350
201 G1YTYrThrAlaAspHisGluArgAlaTyrAlaIleLeuGlnHisSTPTTh 217
|||||
1351 GGGTACACAGCCGACGAGAGCGGCTTACGACCTCTGACGACCTGGAC 1400
217 rIleArgGlyProGlnLysLeuAlaGlnLeuIleSerAlaLeuArg 234
|||||

```

```

1401 CATCCGGGGCCCCGAGGCCAGCTCCGCCAGCTAATTAGCGCCTGGCC 1450
234 InHisArgArgAsnAspValValGlnLysIleLeuGlyLeuMetGlnAsp 250
|||||
1451 AGCACCGGAGAAACGATGTTGTGGAAGAGATTCGTGGCTGATGGAAGAC 1500
251 ThrThrGlnLeuGlnLysThrAspLysLeuAlaLeuProMetSerProSerPr 267
|||||
1501 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1550
267 OLeuSerProSerProIleProSerProAsnAlaLysLeuGlnLysAsn 284
|||||
1551 GCTTAGCCCGAGCCGCAATCCCGACGCCCAACGCGAAACTTGAGAATTCCG 1600
284 ILeuLeuThrValGlnLysProSerProGlnAsp 294
|||||
1601 CTCTCTGACAGCTGTGAGCTTCCCGACAGAT 1632

```

seq_name: gb_pat:AR134339

seq_documentation_block: 2638 bp DNA PAT 16-MAY-2001
LOCUS AR134339
DEFINITION Sequence 22 from patent US 6194151.
ACCESSION AR134339
VERSION AR134339.1 GI:14123244
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2638)
AUTHORS Busfield,S.J.
TITLE Molecules of the TNF receptor superfamily and uses therefor
JOURNAL Patent: US 6194151-A 22 27-FEB-2001;
FEATURES Location/Qualifiers
SOURCE 1.2638
/organism="unknown"

BASE COUNT 625 a 749 c 701 g 563 t
ORIGIN

alignment_scores:
Quality: 1527.00 Length: 294
Ratio: 5.194 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-733-956-2 x AR134339 ..

Align seg 1/1 to: AR134339 from: 1 to: 2638

```

1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValle 17
|||||
1260 ATGACACTCAACAGAAATCCAACTCTTCTGCTGCTTGTAGACCAAGTACT 1309
17 userSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArg 34
|||||
1310 GAGTAGCATCCAGAGAGGACAGTCCCTGACACACACAGCTCAGCAAGG 1359
34 LysGluAspValAsnLysThrLeuProAsnLeuGlnValAlaAsnHis 50
|||||
1360 GGAAGGAGAGCGTGACACAGACCTCCCAACCTTCAGTAGTCAACACC 1409
51 GlnGlnGlyProHisArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||
1410 CAGCAGGCCCCCAGCAGACACATCTGAAAGCTGCTGCTCCATGCA 1459
67 uAlaThrGlyGlyLysSerSerThrProIleLysGlyProLysArg 84
|||||
1460 GGCACACTGGGGGAGAAAGTCCAGACAGCCCATCAAGGGCCCAAGAGG 1509
84 LysProArgGlnAsnLeuHisLysHisPheAspIleAsnGlnHisLeu 100
|||||
1510 GACATCTAGACAGAACCTTACACAGCATTTTGCATCAATGACATTGG 1559

```

	seq_name:	gb_pat:ARI34328	
	seq_documentation_block:		
LOCUS	ARI34328	2612 bp	DNA
DEFINITION	Sequence 3 from patent US 6194151.		PAT
ACCESSION	ARI34328		
VERSION	ARI34328.1	GI:14123233	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2612)		
TITLE	Bustfield,S.J.		
JOURNAL	Molecules of the TNF receptor superfamily and uses therefor		
FEATURES	Patent: US 6194151-A 3 27-FEB-2001;		
	Location/Qualifiers		
source	1..2612		
	/organism="unknown"		
BASE COUNT	661 a	713 c	650 g
ORIGIN		588 t	

Ratio:	5.138	Gaps:	1
Percent Similarity:	95.367	Percent Identity:	94.981
alignment_block:			
US-09-733-956-2	x	AR143328	..
Align seg 1/1	to: AR143328	from: 1	to: 2612

Align seg 1/1 to: AR134328 from: 1 to: 2612

[illegible]

```
alignment_scores:
  quality: 1269.00
```

Length: 259

seq_name: gb_pr:HSJ181J13

1680 GTCATTACCACCGACCTATTGCCCTA 1706

seq_documentation_block:
LOCUS HSJ181J13 118271 bp DNA PRI 21-FEB-2000
DEFINITION Human DNA sequence from clone RPI-181J13 on chromosome 6p12.2-21.1
Contains gene for DR6 (TNFR-related death receptor-6), ESTs, STSs,
GSSs and a Cpg Island, complete sequence.
ACCESSION AL0096801
VERSION AL0096801.18 GI:6478163
KEYWORDS HTG: Cpg Island; DR6; TNFR.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 118271)
AUTHORS Parker A.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Nov 29, 1999 this sequence version replaced gi:6469421.
requests: clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep RPI-181J13 is
from the library RPI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCIPAC2
IMPORTANT: This sequence is not the entire insert of clone
RPI-181J13. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone RPI-181J13 is at 1 in this sequence. The
true left end of clone RPI-1C16 is at 118172 in this sequence. This
sequence was generated from part of bacterial clone contigs of
human chromosome 6, constructed by the Sanger Centre Chromosome 6
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6.
FEATURES
source
1..118271
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="p12.2-21.1"
/clone="RPI-181J13"
/clone_1lb="RPI-1"
4..215
/note="AluSc repeat: matches 93. .303 of consensus"
repeat_region
1668..1735
/note="17 copies 4 mer tcca 91% conserved"
repeat_region
1737..2216
/note="MER31B repeat: matches 1. .485 of consensus"
repeat_region
2223..2350
/note="L2 repeat: matches 2582. .2710 of consensus"
repeat_region
2370..2567
/note="L2 repeat: matches 2244. .2452 of consensus"
misc_feature
complement(2700...3156)
/note="match: GSS: Em:AQ137790"
complement(2712..3043)
/note="match: GSS: Em:AQ094362"

misc_feature
complement(3115..3466)
/note="match: STS: Em:253711
match: STS: Em:HSB3362H9"
repeat_region
3304..3395
/note="46 copies 2 mer tg 73% conserved"
repeat_region
3305..3396
/note="23 copies 4 mer gtgt 73% conserved"
repeat_region
3307..3378
/note="3 copies 24 mer 80% conserved"
misc_feature
3581..4013
/note="match: GSS: Em:AQ030912"
repeat_region
5096..5216
/note="L1P3 repeat: matches 5617. .5742 of consensus"
repeat_region
5394..5677
/note="AluSc repeat: matches 15. .300 of consensus"
repeat_region
5785..6049
/note="L2 repeat: matches 2182. .2455 of consensus"
repeat_region
6052..6085
/note="17 copies 2 mer tt 85% conserved"
repeat_region
6089..7315
/note="L1PB1 repeat: matches 4911. .6152 of consensus"
repeat_region
7326..7384
/note="L2 repeat: matches 2682. .2737 of consensus"
repeat_region
7525..7588
/note="MLTID repeat: matches 1. .65 of consensus"
repeat_region
7767..7902
/note="MLTID repeat: matches 45. .177 of consensus"
repeat_region
7963..8192
/note="MIR repeat: matches 13. .262 of consensus"
repeat_region
8467..8618
/note="MIR repeat: matches 107. .259 of consensus"
repeat_region
8769..8812
/note="MIR repeat: matches 107. .259 of consensus"
repeat_region
8813..9282
/note="L1 repeat: matches 1. .456 of consensus"
repeat_region
9285..9548
/note="L2 repeat: matches 431. .712 of consensus"
misc_feature
9549..9837
/note="match: GSS: Em:AQ041048"
misc_feature
9549..9853
/note="match: GSS: Em:B35323"
repeat_region
9875..10219
/note="L2 repeat: matches 2241. .2624 of consensus"
repeat_region
10464..10534
/note="MER21B repeat: matches 12. .81 of consensus"
repeat_region
10535..10963
/note="MER57A repeat: matches 1. .434 of consensus"
repeat_region
10964..11651
/note="MER21B repeat: matches 81. .778 of consensus"
repeat_region
11647..11717
/note="MER21B repeat: matches 724. .794 of consensus"
repeat_region
12489..18627
/note="L1PA2 repeat: matches 11. .6144 of consensus"
repeat_region
18671..18981
/note="AluYo repeat: matches 1. .310 of consensus"
repeat_region
19040..19169
/note="AluYo repeat: matches 7. .135 of consensus"
repeat_region
19322..21409
/note="L1MA8 repeat: matches 4105. .6230 of consensus"
repeat_region
21410..21717
/note="AluY repeat: matches 1. .308 of consensus"
repeat_region
21718..22040
/note="L1MA8 repeat: matches 3791. .4105 of consensus"
repeat_region
22258..22496
/note="MIR repeat: matches 9. .253 of consensus"
misc_feature
complement(22497..22799)
/note="match: GSS: Em:AQ548159"
22811..23243
/note="match: GSS: Em:AQ711306"
repeat_region
23092..23152
/note="L2 repeat: matches 2688. .2750 of consensus"
repeat_region
23267..23306
/note="L2 repeat: matches 2671. .2709 of consensus"

```

repeat_region 23473..23690
/note="L2 repeat: matches 2064..2298 of consensus"
repeat_region 23900..23957
/note="29 copies 2 mer aa 75% conserved"
repeat_region 24021..24312
/note="AluSp repeat: matches 1..294 of consensus"
repeat_region 24901..25210
/note="AluSg repeat: matches 1..303 of consensus"
repeat_region 25394..25691
/note="AluSx repeat: matches 1..298 of consensus"
repeat_region 26282..26417
/note="L2 repeat: matches 2570..2704 of consensus"
repeat_region 26759..26842
/note="MIR repeat: matches 37..120 of consensus"
repeat_region 27026..27226
/note="MIR repeat: matches 35..249 of consensus"
misc_feature 27372..27965
/note="match: GSS: Em:AQ240568"
repeat_region 27966..28231
/note="L2 repeat: matches 2457..2745 of consensus"
repeat_region 28497..28695
/note="L2 repeat: matches 2041..2227 of consensus"
repeat_region 28784..28974
/note="MIR1 repeat: matches 61..221 of consensus"
repeat_region 29201..29320
/note="5 copies 24 mer 72% conserved"
repeat_region 29289..29324
/note="18 copies 2 mer ct 80% conserved"
repeat_region 29292..29323
/note="8 copies 4 mer tctc 87% conserved"
repeat_region 29325..29610
/note="AluSg1 repeat: matches 1..309 of consensus"
repeat_region 29693..30204
/note="MER9 repeat: matches 1..511 of consensus"
repeat_region 30205..30768
/note="HERV9 repeat: matches 1..568 of consensus"
repeat_region 30770..32408
/note="HERV9 repeat: matches 1113..2772 of consensus"
repeat_region 32416..32572
/note="HERV9 repeat: matches 3711..3867 of consensus"
repeat_region 32659..33272
/note="HERV9 repeat: matches 5303..5916 of consensus"
repeat_region 33273..33776
/note="MER9 repeat: matches 1..503 of consensus"
repeat_region 33964..34050
/note="AluYb8 repeat: matches 231..317 of consensus"
misc_feature 34721..35356
/note="match: GSS: Em:AQ265438"
repeat_region 35982..36176
/note="MER5B repeat: matches 12..175 of consensus"
misc_feature 36177..36582
/note="match: STS: Em:HS124B7T"
misc_feature 36381..37748
/note="CpG island"
evidence-not-experimental
join(36746..37052,59873..60016)

alignment_scores:
Quality: 874.00 Length: 215
Ratio: 4.724 Gaps: 2
Percent Similarity: 86.047 Percent Identity: 80.000

alignment_block:
US-09-733-956-2 x HSJ181J13

Align seg 1/1 to: HSJ181J13 from: 1 to: 118271

1 MetAsnSerThgluSerAsnSerSerAlaSerValArgProLysValLe 17
|||||
62038 AAGCACTCAACAGAAATCAACTCTCTGCTGTAGACCAAGCTACT 62087
17 userSerIleGlnGlyThrValProAspAnThrSerSerAlaArgC 34
|||||

```

```

62088 GAGTACATCCAGGAAGGACAGACTCCCTGACACACAGCTCAGCAAGG 62137
34 1yLysGluAspValaAsnLysThrLeuProAsnLeuGlnValaValaAsnHis 50
|||||
62138 GGAAGGAGAGCGTGAACAGACAGACCTCCAAACCTTCAGGATGACACAC 62187
51 GlnGlnGlyProHisIleArgHisIleLeuLysLeuLeuProSerMetG1 67
|||||
62188 CACCAAGGCCCCACACACAGACATCTCGAACCTCTGCGCCGATGGA 62237
67 uAlaThrGlnGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
|||||
62238 GGGCGAGGCGGAGAGAGTCCAGACAGCCCATCAAGGCGCCCAAGAGG 62287
84 1yHisProArgGlnAsnLeuHisLysHisIlePheAspIleAsnGlnHisLeu 100
|||||
62288 GACATCTTAGACAGAACCTTACACACATTTTACATCAATGAGCATTTG 62337
101 ProTrpMetIleValLeuPheLeuLeuValLeuValaValaVala 117
|||||
62338 CCTGGATGATGTGTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 62387
117 1ySerSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
|||||
62388 GTGCACTATCCGGAAGAGCTCGAGACTTGAGAAAGGGGCGCCGACAG 62437
134 sProSerAlaIleValaGluLysAlaGlyLeuLysLysSerMetThrPro 150
|||||
62438 ATCCAGATGCCATTGTGGAAGAGCGAGGCTGAAGAAATCATGACTCCA 62487
151 ThrGlnAsnArgGluLysTrpIleTyrrYrcYasGlnGlyHisIleLeu 167
|||||
62488 ACCCAAGACCGGAGGAATGATGATCTGCAATGCGCATGTGAGGCC 62537
167 pIleLeuLysLeuValaIAlaAlaGlnVal 176
|||||
62538 TTCTCTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 62587
177 1ySerGlnTrpLysAspIleTyrrGlnPheLeuCys 188
|||||
62588 CCCTGTTTCTGAAGGAATGCGAATGCGTGTATTGAGAAATCCCTTTG 62637
189 AsnAlaSer 195
|||||
62638 AGACCAACACACTGATCTACTGCTGTATTGAAAGAGATT 62682

seq_name: gb_htg:AL512350

seq_documentation_block:
LOCUS AL512350 150180 bp DNA 10-JUL-2001
DEFINITION Homo sapiens chromosome 6 clone RP11-150D3, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
ACCESSION AL512350
VERSION AL512350.7 GI:13169571
KEYWORDS HTG: HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 150180)
AUTHORS Cordy,N.
TITLE Direct Submission
SUBMITTED (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
REQUESTS: clones requested at: sanger.ac.uk
On Feb 26, 2001 this sequence version replaced gi:13161707.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bai50D3

```


JOURNAL

COMMENT

Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 1, 2000 this sequence version replaced g1:7272272.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L8128

Center Clone name: 304_O_4

* NOTE: This record contains 151 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 765: contig of 765 bp in length
* 766 865: gap of 100 bp
* 866 1613: contig of 748 bp in length
* 1614 1713: gap of 100 bp
* 1714 2478: contig of 765 bp in length
* 2479 2578: gap of 100 bp
* 2579 3318: contig of 740 bp in length
* 3319 3418: gap of 100 bp
* 3419 4158: contig of 740 bp in length
* 4159 4258: gap of 100 bp
* 4259 5016: contig of 758 bp in length
* 5017 5116: gap of 100 bp
* 5117 5898: contig of 782 bp in length
* 5899 5998: gap of 100 bp
* 5999 6729: contig of 731 bp in length
* 6730 6829: gap of 100 bp
* 6830 7607: contig of 778 bp in length
* 7608 7707: gap of 100 bp
* 7708 8462: contig of 755 bp in length
* 8463 8562: gap of 100 bp
* 8563 9304: contig of 742 bp in length
* 9305 9404: gap of 100 bp
* 9405 10153: contig of 749 bp in length
* 10154 10253: gap of 100 bp
* 10254 11009: contig of 756 bp in length
* 11010 11109: gap of 100 bp
* 11110 11845: contig of 736 bp in length
* 11846 11945: gap of 100 bp
* 11946 12729: contig of 784 bp in length
* 12730 12829: gap of 100 bp
* 12830 13572: contig of 743 bp in length
* 13573 13672: gap of 100 bp
* 13673 14397: contig of 725 bp in length
* 14398 14497: gap of 100 bp
* 14498 15240: contig of 743 bp in length
* 15241 15340: gap of 100 bp
* 15341 16086: contig of 746 bp in length
* 16087 16186: gap of 100 bp
* 16187 16907: contig of 721 bp in length
* 16908 17007: gap of 100 bp
* 17008 17806: contig of 799 bp in length
* 17807 17906: gap of 100 bp
* 17907 18678: contig of 772 bp in length
* 18679 18778: gap of 100 bp
* 18779 19523: contig of 745 bp in length
* 19524 19623: gap of 100 bp
* 19624 20368: contig of 745 bp in length

* 20369 20468: gap of 100 bp
* 20469 21217: contig of 749 bp in length
* 21218 21317: gap of 100 bp
* 21318 22052: contig of 735 bp in length
* 22053 22152: gap of 100 bp
* 22153 22920: contig of 768 bp in length
* 22921 23020: gap of 100 bp
* 23021 23776: contig of 756 bp in length
* 23777 23876: gap of 100 bp
* 23877 24606: contig of 730 bp in length
* 24607 24706: gap of 100 bp
* 24707 25456: contig of 750 bp in length
* 25457 25556: gap of 100 bp
* 25557 26333: contig of 777 bp in length
* 26334 26433: gap of 100 bp
* 26434 27207: contig of 774 bp in length
* 27208 27307: gap of 100 bp
* 27308 28037: contig of 730 bp in length
* 28038 28137: gap of 100 bp
* 28138 28919: contig of 782 bp in length
* 28920 29019: gap of 100 bp
* 29020 29771: contig of 752 bp in length
* 29772 29871: gap of 100 bp
* 29872 30649: contig of 778 bp in length
* 30650 30749: gap of 100 bp
* 30750 31545: contig of 796 bp in length
* 31546 31645: gap of 100 bp
* 31646 33375: contig of 730 bp in length
* 33376 33475: gap of 100 bp
* 33476 33523: contig of 762 bp in length
* 33528 33337: gap of 100 bp
* 33338 34068: contig of 731 bp in length
* 34069 34168: gap of 100 bp
* 34169 34911: contig of 743 bp in length
* 34912 35011: gap of 100 bp
* 35012 35741: contig of 730 bp in length
* 35742 35841: gap of 100 bp
* 35842 36621: contig of 780 bp in length
* 36622 36721: gap of 100 bp
* 36722 37482: contig of 761 bp in length
* 37483 37583: gap of 100 bp
* 37583 38332: contig of 750 bp in length
* 38333 38432: gap of 100 bp
* 38433 39210: contig of 778 bp in length
* 39211 39310: gap of 100 bp
* 39311 40095: contig of 785 bp in length
* 40096 40195: gap of 100 bp
* 40196 40978: contig of 783 bp in length
* 40979 41078: gap of 100 bp
* 41079 41831: contig of 753 bp in length
* 41832 41931: gap of 100 bp
* 41932 42673: contig of 742 bp in length
* 42674 42773: gap of 100 bp
* 42774 43516: contig of 743 bp in length
* 43517 43616: gap of 100 bp
* 43617 44358: contig of 742 bp in length
* 44359 44458: gap of 100 bp
* 44459 45217: contig of 759 bp in length
* 45218 45317: gap of 100 bp
* 45318 46067: contig of 750 bp in length
* 46068 46167: gap of 100 bp
* 46168 46907: contig of 740 bp in length
* 46908 47007: gap of 100 bp
* 47008 47735: contig of 728 bp in length
* 47736 47835: gap of 100 bp
* 47836 48605: contig of 770 bp in length
* 48606 48705: gap of 100 bp
* 48706 49485: contig of 780 bp in length
* 49486 49585: gap of 100 bp
* 49586 50330: contig of 745 bp in length
* 50331 50430: gap of 100 bp
* 50431 51176: contig of 746 bp in length
* 51177 51276: gap of 100 bp

```

*      51277      52010: contig of 734 bp in length
*      52011      52110: gap of 100 bp
*      52111      52868: contig of 758 bp in length
*      52869      52968: gap of 100 bp
*      52969      53704: contig of 736 bp in length
*      53705      53804: gap of 100 bp
*      53805      54496: contig of 692 bp in length
*      54497      54596: gap of 100 bp
*      54597      55321: contig of 725 bp in length
*      55322      55421: gap of 100 bp
*      55422      56136: contig of 715 bp in length
*      56137      56236: gap of 100 bp
*      56237      56945: contig of 709 bp in length
*      56946      57045: gap of 100 bp
*      57046      57752: contig of 707 bp in length
*      57753      57852: gap of 100 bp
*      57853      58556: contig of 704 bp in length

```

alignment_scores:

```

Quality: 408.00      Length: 127
Ratio: 4.295      Gaps: 2
Percent Similarity: 74.803      Percent Identity: 64.567

```

alignment_block:

US-09-733-956-2 x AC026260/rev ..

Align seg 1/1 to reverse of: AC026260 from: 1 to: 125419

```

89 AsnLeuHisLysHisPheAspIleAsnGluHisLeuProTrpMetIleVal 105
||| ||||||||||||||||||||||||||||||||||||||||||||
69948 AACCTTACACAGCATTTTGACATCAATGACGATTTGCCCTGATGATTGT 69899

105 lleuPheleuLeuValleuValIleValIleValCysSerIleArgL 122
||||||||||||||||||||||||||||||||||||||||||
69898 GCTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 69849

122 ysserSerArgThrleuLysLysGlyProArgGlnAspProSerAlaIle 138
||||||||||||||||||||||||||||||||||||||||||
69848 AAAGCTCGAGGACTCTGAAGAGGCGCCGCGCAGATCCCATGCCATT 69799

139 ValGluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgL 155
||||||||||||||||||||||||||||||||||||||||||
69798 GTGGAAGAAGCGAGGCTGAAGAAATCCATGATGATCAACCAACCGGGA 69749

155 uLysTrpIleTyrtYrCysAsnGlyHisGlyIleAspIleLeuLysLeuV 172
||||||||||||||||||||||||||||||||||||||||||
69748 GAATGATCTACTACTGCAATGGCCATGGTGAACCTTCTCTCTCTCTC 69699

172 alaIalaIleVal..... 176
|||||
69698 CTGCTCAGAGATCTCTCTACCACTTTTCCCTAGTCCCTGTTTCTGAA 69649

177 GlySerClnTrpLysAspIleTyrglnPheLeuCysAsnAlaser..... 191
|||
69648 GGAATGGGAATGGCTAGTTGTGGAATCTCTTTGTAGACCAAAACACTAC 69599

192 .....GluArgGluVal 195
|||||
69598 TTGATCTAGTCGTGTTATAGAAAAAGAGATT 69568

```

THIS PAGE BLANK (USPTO)